THREE ESSAYS ON CONTEMPORARY ISSUES IN CLIMATE CHANGE AND URBAN DEVELOPMENT

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sanket Roy

August 2017

THREE ESSAYS ON CONTEMPORARY ISSUES IN CLIMATE CHANGE AND URBAN DEVELOPMENT

Sanket Roy, Ph.D.
Cornell University 2017

The first paper contributes to the existing literature which looks at the effects of rising temperatures on energy demand, more specifically household energy consumption. We use a novel methodology to isolate extensive margin adjustments made in response to a changing climate from the intensive margin adjustments made in response to changes in weather. By controlling for both the contemporaneous temperature distribution as well as the lagged moving average, we are gaining on two fronts. Firstly, we are able to get more precise estimates of the effect of short-run changes in weather on residential energy consumption. Secondly, by comparing the two effects, we are able to back out the extensive margin effects on residential energy consumption made in response to changes in longer term climate. Our estimates imply significant impacts for both the intensive and the extensive margins and a U-shaped response function of residential energy consumption with respect to contemporaneous temperature.

Rising income inequality is a cause of concern for policy makers in many developed and developing countries. Another concern in several developing countries is the rise in informal settlements. This second paper attempts to connect the two and see if rising income inequality contributes to the problem of rising informal settlements. We model informal settlements explicitly and show that having more than two income groups matter in our set up. We find that rising income inequality does contribute to the welfare of the poor in a negative

way and also leads to a rise in number of informal settlers, hence contributing to the existing problem. We extend the model further to incorporate tax on housing in the formal sector and tax revenue going into the provision of public utilities in the informal sector and look at the effects of rising income inequality. We also look at movements of households across sectors and find interesting implications of these to the city equilibrium.

The third paper uses the monocentric city model set up and enhances the model to include for different types of housing and differences in their respective construction costs. We show the importance of these features by comparing the equilibrium to a situation where these features were not there. We find interesting implications of the new features of the model on the city equilibrium. We show that having the highest bid-rent does not ensure a household a house in that place. We also do some comparative static analysis and show how changes in construction costs changes the equilibrium of the city.

BIOGRAPHICAL SKETCH

Sanket Roy was born in Durgapur, a small city in West Bengal, India, in December 1988. He finished school till the tenth standard at St. Xavier's, Durgapur and then went to Bidhan Chandra Institution, Durgapur for completing the plus two level. After finishing school, Sanket moved to Kolkata for his undergraduate studies at Jadavpur University. He successfully completed Bachelor of Arts (Honors) in Economics from Jadavpur University in 2010. For enhancing his skills further, he joined Indian Statistical Institute, Delhi and earned the degree of Master of Science and Quantitative Economics, in 2012, before coming to Cornell University for graduate studies.

To *Maa* and *Baba*

ACKNOWLEDGEMENTS

I would like to thank all my advisors, Professor Ravi Kanbur, Professor Antonio Bento and Professor Shanjun Li for their guidance and support throughout these years. I am especially thankful to Professor Ravi Kanbur for all his help and advice at times when the going got tough. Professor Kanbur really played a pivotal role in completion of my PhD and I cannot thank him enough for that. I have really learned a lot from all of my advisors and they have indeed enriched my Cornell experience.

I would like to thank Suvro Chatterjee who taught me when I was young and helped immensely in both my academic and non-academic development. It was only because of his teachings and influence that I chose to study Economics without the support of my family. I would be eternally grateful to *Sir* for motivating me to do what I wanted to do and guiding me through various stages of life.

I am really blessed to have two wonderful human beings as my parents. I would like to thank Maa and Baba from the bottom of my heart for taking care of me throughout my life and fulfilling my needs, even at times when it was tough for them to do so. I feel really lucky to have them in my life and both of them have contributed in numerous ways to make sure that I reach my goals.

Life becomes easy when one finds a friend in his sister and becomes even easier when one is blessed to find a brother in his brother-in-law. The presence of Didi and Suvro Da are the primary reason for me not worrying about my parents even when staying several thousand miles apart. I would like to thank Didi and Suvro Da and also Jethu and Jethima for being by the side of my parents always and helping me having a peaceful PhD life.

Didi and Suvro Da did another amazing thing by presenting us the best gift ever in the form of my beloved nephew, *Paglu*. Perhaps *Paglu* is too young to understand, but he has immense contribution in my PhD life. Throughout these five years, I missed him a lot and it were the numerous stories and videos of him growing up, starting to walk, then growing up more to increase the level of *bandramo*, which helped me forget everything else and have a good laugh at the end of the day.

I am thankful to Professor Saikat Sinha Roy for being a true friend, philosopher and guide for all these years. I do share a special relationship with him and it is his guidance and support right from my undergraduate days which helped me in getting through to ISI, Delhi for doing masters and then coming to Cornell for my PhD. I am eternally grateful to him and Thakuma for helping my family in Kolkata when I was far away here in Ithaca.

I met Mehreen ten years back during my undergraduate days and within a few days we struck a chord. Throughout these years, I have been extremely fortunate to have her by my side always. I cannot thank her enough for the countless number of times she has bailed me out of difficult circumstances. I will be eternally indebted to her for her support and guidance throughout the last ten years and especially for the pivotal role that she played in the last five years. Without her presence in Ithaca, I do not think I would have crossed the line. I am thankful to her for being my best friend till now and also choosing to be my best friend for life. I am also thankful to her parents, Kaku and Kakima, for being the house doctor for us throughout these years and providing support in times of need.

I would also like to thank several people in Cornell who made my stay in Ithaca so enjoyable. The *Bengali Group* made my life so much easier and it was only because of them that a foody like me survived in Ithaca. I am thankful to

Mota aka Anirudra for being the catalyst of all the fun and we really missed him when he moved out of Ithaca. I also found a big brother cum friend in *Master* aka Gourav and would like to thank him for helping me throughout in various capacities. I am also thankful to Siddy, Nidhi, Poornima and the *Villager Group* for supporting me throughout these years. Survival in Ithaca would have been a lot more difficult without the *Bengali Group* and the *Villager Group*. I am also thankful to Debi Da, Tirupam, Jyoti and Isha for all of their help and support throughout out these years.

TABLE OF CONTENTS

CHAPTER 1

THE SHORT AND LONG RUN IMPACTS OF TEMPERATURE ON U.S. RESIDENTIAL ENERGY CONSUMPTION

1.1 Introduction

There is a substantial literature which has looked at the effects of climate change on energy consumption or domestic fuel use. The broad intuition is that due to climate change the summer months are becoming hotter and in the winter, there is less cold. Therefore intuitively, one should see a higher consumption of energy in the summer and a lower consumption of energy during winter. It is important to note here and distinguish the fact that higher temperatures affect both the short run and the long run behavior of the households.

The motivation for doing this work is the fact that GHG emissions related to energy use account 70% of total global emissions. This includes consumption by households as well as industries and our focus in this paper will be household consumption. The common finding is that in the short run households respond by using more energy. Climate change will affect energy usage by changing how agents respond along both intensive and extensive margins of adjustment. For example, in the short run, they can adjust by using existing equipment more intensively whereas in response to longer run climatic changes, they can potentially adjust by buying more equipment such as air conditioners. Therefore, just controlling for the contemporaneous temperature would give us two effects, change in residential energy use due to change in weather (weather is the deviation of today's temperature from the long-term trend) and change in energy consumption due to change in climate (long-term trend of temperature).

Moreover climate impacts our response in two ways, a $1^{\circ}F$ rise in climate trend would lead to an increase in today's temperature which would induce households to change their energy consumption and it can also affect households decision in adjustments in the extensive margin.[1] In this paper I use historical weather data, to estimate the impacts of both contemporaneous weather shocks as well as long term climate trends on average state-level energy usage for the US.

In this paper, we use a novel methodology, inspired from the seminal work of Solon (1992). We create moving averages for each temperature bin and also control for that in our specification. It captures households reaction to the trends that they observe, for example a household might choose to buy a new air conditioner in response to the trend that it sees which would increase consumption of energy. Controlling for moving average of temperatures also gives us more precise estimates of the contemporaneous effect because if not controlled, the contemporaneous effect would have been a combined effect of contemporaneous change in temperature as well as adaptation to climate change. Dell, Jones, and Olken (2009, 2012, 2014) also does something similar where they control for lagged temperatures in their specification in their study of climate change and income.

We find a U shaped response function for residential energy consumption as a function of temperatures. Interestingly, we find significant impacts for one more day in the extreme bins on residential energy consumption. Controlling for climate, we find significant evidence of adaptation to climate change and this also allows us to get better estimates of the effects of changes in weather on residential energy consumption. We also find significant variation of responses

[1] In this paper we create bins for temperature and use that as a control.

across climatic regions across the US.

1.2 Literature Review

There are many papers till date which have tried to establish the causal link between climate and energy consumption and supply but the literatures primary focus has been on the demand or consumption side. Moreover, even on the demand side, most of the papers have looked at the residential sector. The empirical studies on this issue can be broadly divided into two parts- one is the study about the intensive margin or the short-run and the other is studying about the extensive margin or the long-run.

Vaage (2000) and Mansur etal. (2008) build discrete choice model of households fuel choice and uses cross-sectional data for the analysis. Vague uses Norwegian Central Bureau of Statistics survey of 2289 households' energy consumption to look at heating technology and energy use in Norway and find that warmer households spend 30 percent less on fuel. Mansur etal. do a similar analysis for USA by matching weather data to households and firms survey data to create a cross-sectional data for their study. Their paper conclude that firms and households will consume more electricity and oil due to warmer summer but households will reduce consumption of natural gas as a result of warmer winters. The primary drawback of using cross-sectional data is that of omitted variable bias because one cannot control for unobserved heterogeneity across households and firms which can be potentially correlated to temperatures or precipitation.

Another strand of the literature uses univariate time series data to do the analysis. Franco and Sanstad use data reported by the California Independent System Operator for 2004, where they explain the time series variation in hourly load in electricity over a course of one year and regress them on average daily temperature. They find a non-linear impact on electricity load but a linear impact on maximum temperature on peak demand. Mostly similar but slightly different approach was uses by Crowley and Joutz (2003), where they look at the effects of temperature on electricity load. They use hourly data for Pennsylvania, New Jersey and Maryland Interconnection and find that a $3.6°F$ rise in temperature would result in 3-8% rise in consumption of electricity.

Auffhammer and Aroonruengsawat (2011, 2012b) and Deschenes and Greenstone (2011) are recent papers where they use panel data to study the problem. The advantage of using panel data is the fact that it allows the econometrician to control for heterogeneity in unobservables. Deschenes and Greenstone use panel data at the state level in the US for 1963-2002 and conclude that with "business as usual" residential energy consumption will increase by 11% by 2099.[2] On the other hand Auffhammer and Aroonruengsawat (2011, 2012b) uses panel data at the household level and uses information on electricity billings of households in California for the years 2003-2006. Eskeland and Mideksa (2010) study European countries' annual electricity demand and find very small effects of temperature on electricity consumption.

Moving on to the extensive margin, it should be noted that our knowledge is much limited in comparison to the intensive margin. Rapson (2011) builds a structural discrete time, infinite horizon model where households choose whether to but a unit of a durable good or not in each period. The author goes

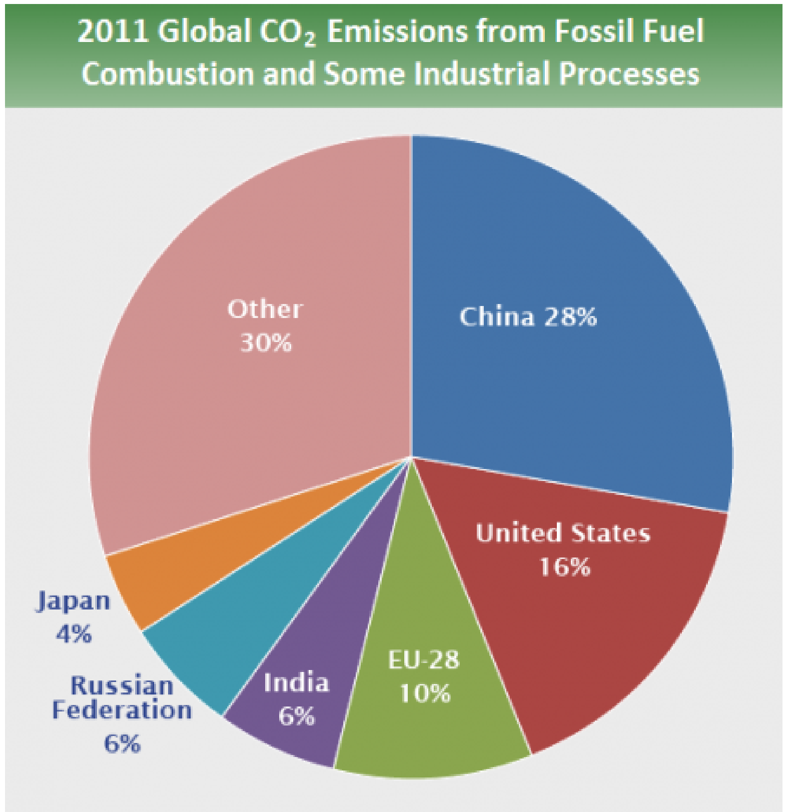
[2]Deschenes and Greenstone (2011) use climate prediction models to create bins of future temperatures.

on to estimate unit demand elasticity where results show elasticity for central air conditioner range between 0.7-1 and for room air conditioner range between 0.2-0.3. Akpinar-Ferrand and Singh (2010) study demand for air conditioning in India. They predict that with a growing economy and higher incomes energy demand will increase significantly due to air conditioning usage. They predict that by 2100, with a 3.7$^{\circ}C$ increase in temperature, there would be an additional demand of energy in the range of 750,000 GWh to 1,350,000 GWh.

This paper follows in line with this extensive literature which tries to understand the different dimensions in which changing climate change human behavior. Our novel approach allows us to get precise estimates for both the intensive and extensive margin adjustments made by households across the US. With the estimates in place, one can further take the route taken by Deschenes and Greenstone (2011) to get precise monetary valuations of predicted change in residential energy consumption using the climate prediction models.

1.3 Climate Change and Energy Consumption

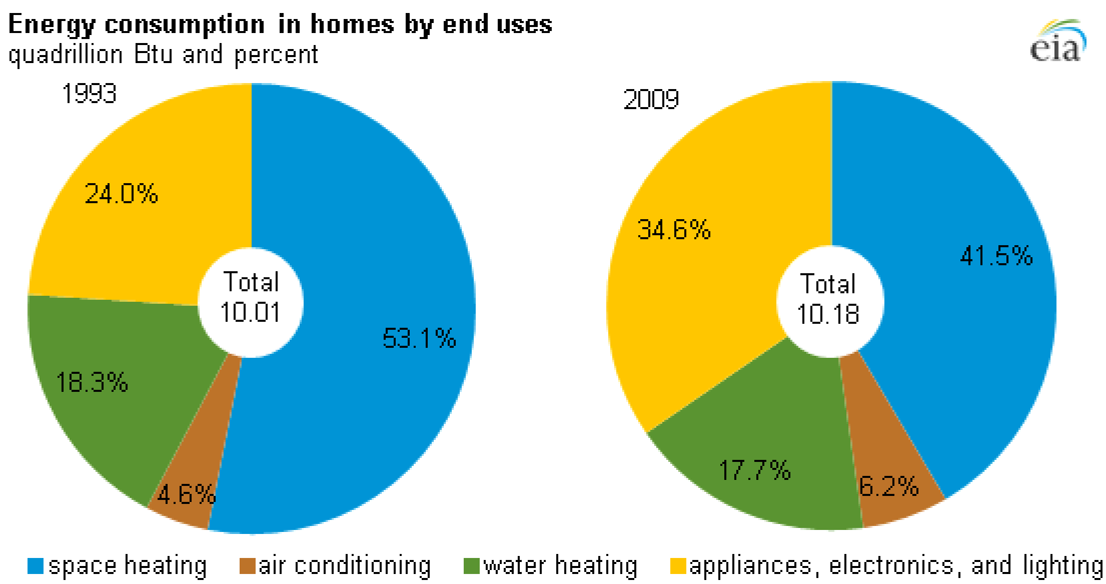
Climate change is a central topic of interest in policy debates these days. Importantly, the causes and consequences of global warming are not limited to a particular area or nation, which makes it even more complicated to tackle. IPCC (Intergovernmental Panel on Climate Change), 2014 reports that almost certainly 95% of global warming is caused due to human activities. They predict that during the 21st century, if one considers the lowest emissions model, surface temperatures are likely to rise by $0.5-3.1^{\circ}F$ and it might go up to a range of $4.7-8.6^{\circ}F$ if one considers the highest emissions model. Coal, petroleum and natural gas are our main sources of energy in present times. It is estimated that around 86% of total energy used in 2005 was due to fossil fuel combustion and in the US it was about 85%. Fossil fuels also contribute to climate change since they contribute to around 80% of greenhouse gas emissions and 98% to CO_2 emissions. Natural gas, which is one of the biggest energy sources for residential energy consumption is also considered to be a huge contributor to global warming because of the tremendous energy required to liquefy and transport it to different places. Figure 1.1 shows the major contributors to CO_2 emissions from fossil fuel and one can see that the US and China are the highest contributors. Therefore, in the situation we are in, it is important to study the behavior of economic agents and how they respond to climate change.



Source: IPCC Report, 2014.

Figure 1.1: Major Contributors to CO_2 Emissions

In the US, residential sector is one of the major users of energy and figure 1.2 shows the energy consumption of households by different end-uses for the years 1993 and 2009. The figure shows that proportion of total energy used has decreased for space heating but has increased for cooling and appliances, electronics and lighting. It is true that over the years due to changes in technology, housing types, there has been decrease in energy consumption per capita but total energy consumption by the residential sector has been increasing. Figure 1.3 from the US Energy Information Administration shows that there is steady increase in residential energy consumption except for 2016 and that is indeed a substantial portion of the total energy consumption. For the last year in our sample, that is 2013, residential energy consumption accounts for approximately 22% of the total consumption of energy in primary use. Figure 1.4 shows the total residential energy used from 1980 – 2009 for the four broad climatic regions in the US which shows that over this time period, residential energy consumption has been increasing more for the West and South in comparison to the other two climatic regions.



Source: U.S. Energy Information Administration

Figure 1.2: Residential Energy Consumption by End Uses (1993 vs 2009)

Energy consumption estimates by sector

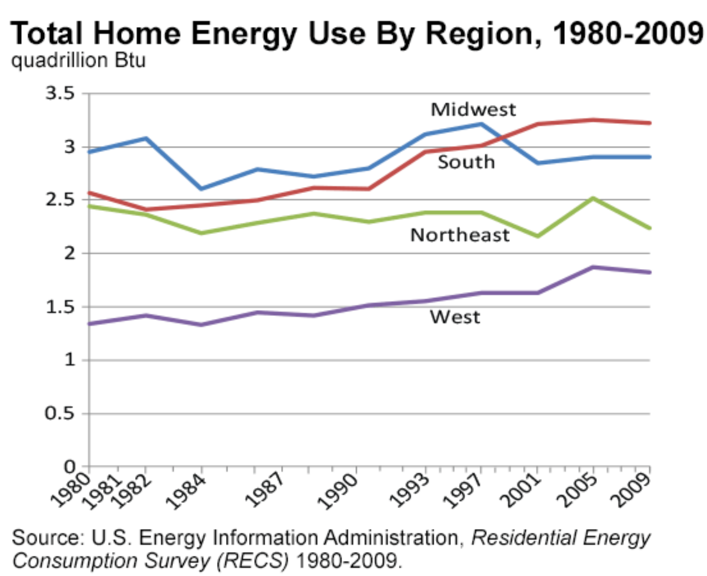
trillion Btu

January to May	2016	2015	2014	2013	2012
End-Use Sector					
Residential	8,751	9,508	9,923	9,308	8,352
Commercial	7,451	7,740	7,853	7,511	7,175
Industrial	12,697	12,938	13,024	12,861	12,814
Transportation	11,504	11,192	10,944	10,803	10,783
Primary Total	40,392	41,375	41,757	40,473	39,109

Source: U.S. Energy Information Administration, Monthly Energy Review – Table 2.1

Source: U.S. Energy Information Administration

Figure 1.3: Energy Consumption by Sector (2012 to 2016)

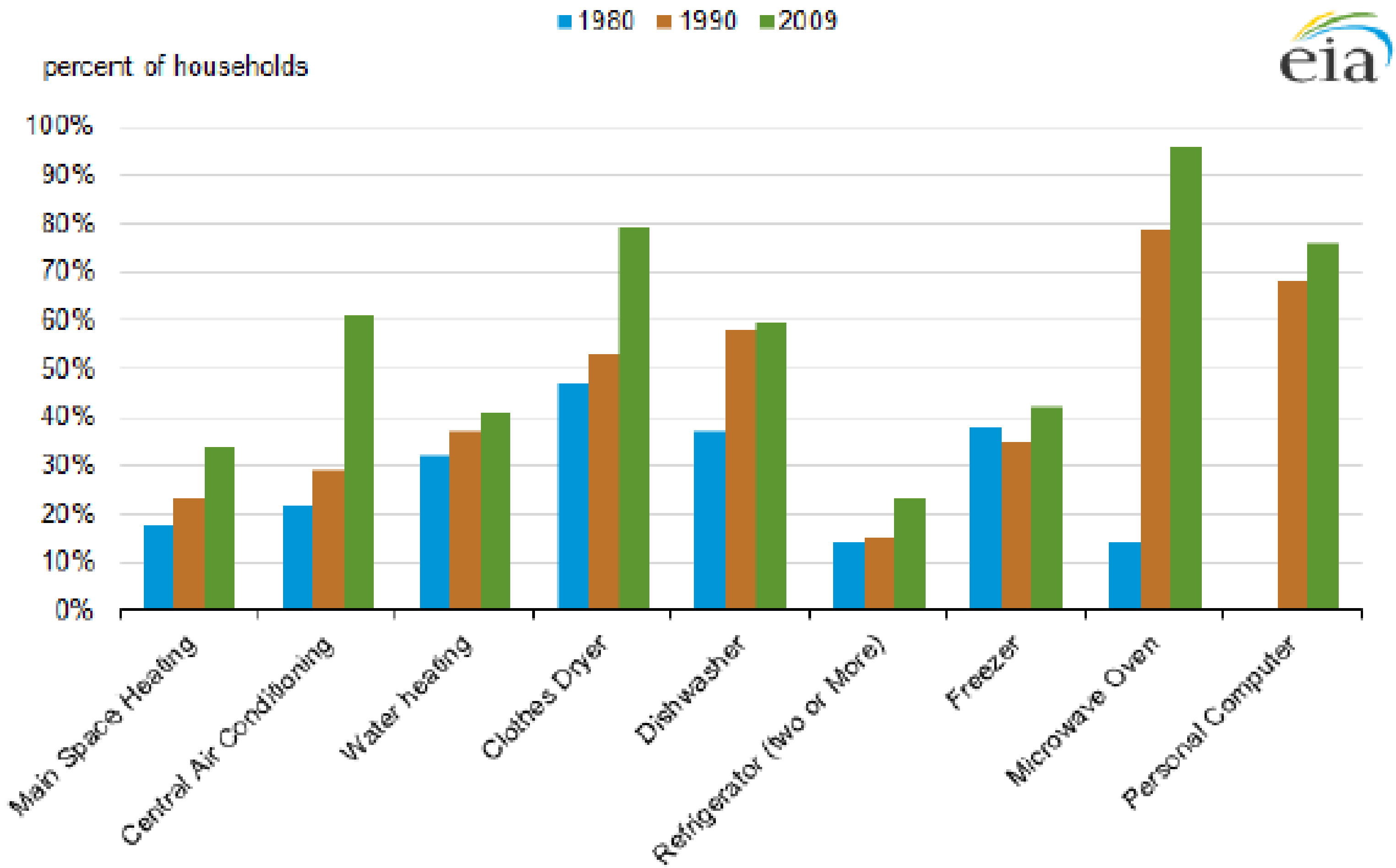


Source: U.S. Energy Information Administration

Figure 1.4:

Figure 1.5 from the US Energy Information Administration gives us some idea about the usage of air conditioners and space heating over the period of 1980 – 2009. We can clearly see that there has been substantial increase in usage of air conditioner and space heating and therefore this should result in higher residential energy consumption.

A better understanding of the behavior of households in necessary for sustainable development. This paper contributes in that respect, since a better understanding of household response to changing temperatures would allow us to predict the future impact in a more precise manner. The period under study has coincided with steady penetration of heating and cooling facilities in the US. Therefore, this might also help in understanding the how demand of residential energy will change in other countries where cooling and heating facilities have not yet penetrated.



Sources: Hojjati and Wade (2012) and U.S. Energy Information Administration, Residential Energy Consumption Surveys, 1980, 1990, and 2009.

Figure 1.5: Access to Energy by End Uses (Percentage of Households)

1.4 Data and Summary Statistics

In this analysis we have used detailed data on climate variables,more specifically temperature and precipitation and data on household energy consumption and state level data on population and GDP. In what follows, we discuss the sources of the data and how they have been aggregated at the state level in order for us to do the analysis.

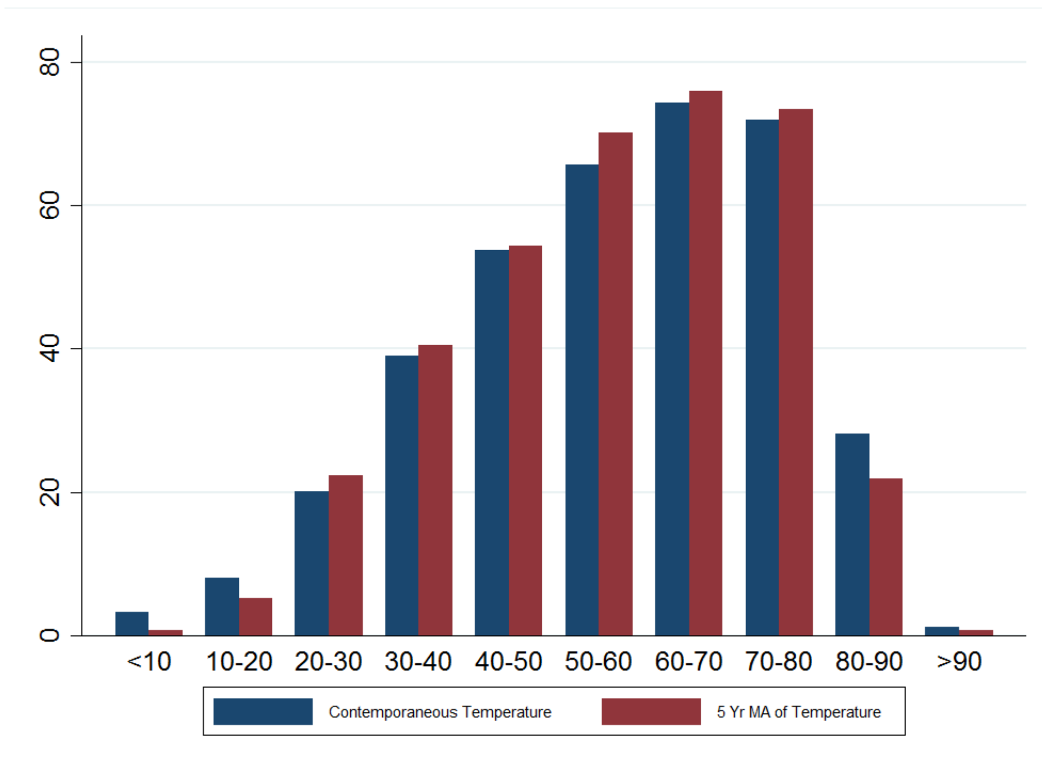
Climate Data: The data on daily temperatures and rainfall has been collected from the National Climatic Data Centers Cooperative Station Data (NOAA 2008) for 1965-2013. This dataset provides detailed information on daily temperatures and precipitation which is collected using 20000 weather stations spread across the entire US. The key variables for us are the daily maximum and minimum temperatures and the daily precipitation and we have used the daily maximum and minimum temperatures to construct the daily mean temperature at the weather station level. Then we match stations to counties for each year and use the weather stations within 200 kms radius of each county centroid. [3] We compute the average weather across all weather stations within 200 kms radius from the county centroid and use the inverse distance square as weights to get all daily weather data at the county level. In order to aggregate to the county-year level, we compute the total yearly rainfall in a county and then create indicator variables based on annual rainfall in a county. Each indicator variable corresponds to a 10-inch bin of annual total rainfall and we have 7 bins ranging from below 10 inches to above 60 inches. For the data on temperature we first create bins starting from less than 10°F to greater than 90°F, with 10°F intervals in between. Then, using these, we calculate the number of days in a year that a

[3]We have also considered radius of 100, 150 and 250 kms, the results are in the robustness section.

county has experienced temperatures in each of the 10 bins. Finally, we take average, weighted by the population in each county, to aggregate all our relevant variables to the state-year level.

We also construct moving average of the average temperature for each day for every county and follow the same process as described above to construct moving average of the climate variables for each state for the years 1975-2013. The construction of the moving averages gives us the long-term climatic trends for every state. We try to separate climatic trends to climate shocks to get at adaptation to climate change. More specifically, controlling for climatic trends/weather would allow us to isolate the intensive margin and extensive margin adjustment is response to changing climate. A more detailed explanation is in the next section. Figure 1.6 compares the distribution of contemporaneous temperature to the distribution of the 5 years moving average of mean temperatures that have been constructed by the method discussed above. We can see from the figure that there are signs of more extreme temperatures than before, especially for the two extreme bins where we can see that the average number of days in the lowest bin is higher for contemporaneous temperature and the opposite is true for the other extreme. We can also see that except for the extreme two temperature bins on either side of the distribution, the number of days in the bins from $20^{\circ} - 80^{\circ}F$ are more for 5 years moving average in comparison to the contemporaneous temperatures.

Energy Data: We have taken the energy data from the U.S. Energy Information Administration (EIA) State Energy Data System and we are using data from 1975-2013. The comprehensive data source has information on energy consumption by households, industrial sector and transportation services. In



Notes: This figure represents the temperature distributions of the contemporaneous and moving average variables. Each bar represents the average number of days in each temperature bin. Averages are constructed across all states and years in sample.

Figure 1.6: Comparison: Distributions of the Contemporaneous and 5 Yr Moving Average Temperature

this paper, we focus on the residential energy consumption which is recorded in British Thermal Units (BTU) and the data also includes state level energy prices, state level population and GDP. The data used as consumption of energy by residents of a state is an aggregate of primary energy use and the loss in energy due to transmission and distribution of energy. As has been noted in Greenstone and Deschenes (2011), this loss in energy due to transmission etc. accounts for about two-thirds of total energy consumption. Given, the fact that climate data is available at the station level, we could have potentially done the analysis at the county level and this would have saved us from potential errors of aggregating at the state level but the unavailability of detailed data on energy consumption the county level leaves us with no other choice than to do it at the state level.

Table 1.1 presents the summary statistics for our variables of interest for the entire US and across 9 different climatic regions. The first four columns present information on the residential energy consumption and the last four columns present information on average daily temperature. As expected, we see significant variation in energy consumption and average temperatures across the US. West, South and Southwest have the highest three mean temperatures in our data and its effect on residential energy consumption is also reflected by the fact that their energy consumption is on the higher side. Table 1.2 given summary statistics for the extreme temperature bins in our paper. We see that average number of days in the contemporaneous temperatures bins is higher than the average number of days in the corresponding moving average bins.

Figures 1.7 and 1.8 gives us some idea about the correlation between residential energy consumption and the climate. Figure 1.7 uses 5 years moving

Table 1.1: Summary Statistics: Residential Energy Consumption and Average Daily Temperature

	Residential Energy Consumption (Quad BTU)				*Average Daily Temperature (F)*			
	Mean	SD	Min	Max	Mean	SD	Min	Max
All States	0.38	0.34	0.03	1.74	53.89	7.49	37.03	74.21
By Climate Regions								
Ohio Valley	0.53	0.28	0.12	1.03	53.75	2.90	48.30	60.50
Upper Midwest	0.44	0.20	0.20	0.87	46.72	2.41	40.62	52.53
Northeast	0.36	0.35	0.03	1.25	50.12	3.94	42.00	58.08
Northwest	0.25	0.14	0.07	0.52	49.65	2.03	43.31	53.89
South	0.41	0.41	0.15	1.74	62.58	4.29	53.28	69.76
Southeast	0.52	0.26	0.17	1.28	62.94	4.84	56.08	74.21
Southwest	0.18	0.10	0.06	0.40	55.15	7.97	45.57	69.60
West	0.74	0.64	0.05	1.52	61.71	1.70	57.16	64.67
Rockies	0.07	0.04	0.03	0.17	45.26	3.41	37.03	53.93

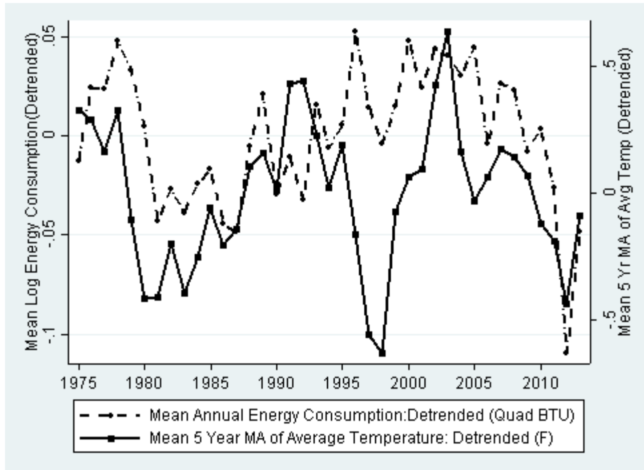
Notes: This table has been constructed by using 1872 state-year observations over the years 1975-2013. The numbers reported are averages across all states and years, or for all states within a climate region.

Table 1.2: Summary Statistics: Number of Days that Contemporaneous and 5 Year Moving Average Temperature is in Extreme Bins

		Mean	Standard Deviation	Maximum
Number of Days $TMean^{Cont}$ is	Below 10 F	5.83	10.16	78.81
	10-20 F	11.16	11.13	49.30
	80-90 F	21.51	27.25	130.38
	Above 90 F	1.31	5.84	48.51
Number of Days $TMean^{MA}$ is	Below 10 F	1.83	6.20	51.39
	10-20 F	9.44	15.00	68.71
	80-90 F	15.64	27.10	120.73
	Above 90 F	0.90	4.82	44.61

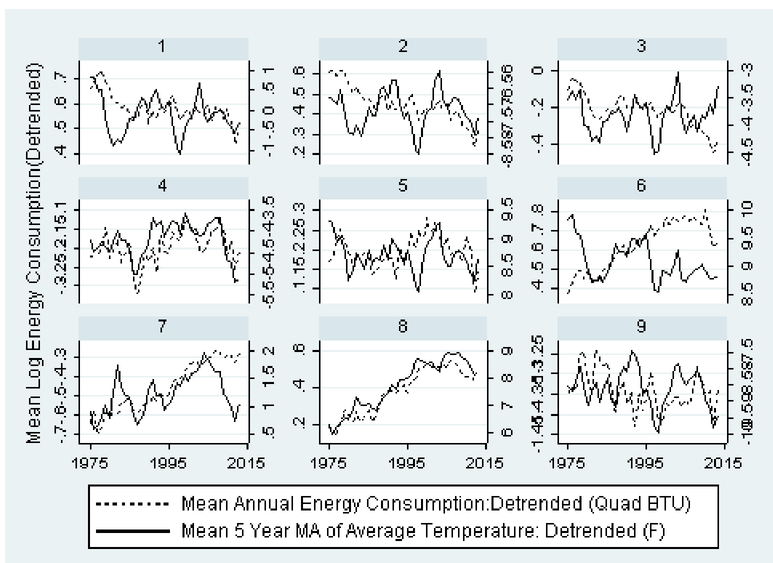
Notes: This table has been constructed by using 1872 state-year observations over the years 1975-2013. The numbers reported are capture the average number of days that the current year's average temperature is in the four extreme temperature bins. The bottom panel reports the same for the past five years' moving average of temperature.

average as a notion of climate and Figure 1.8 uses 10 years moving average as a measure of climate. Both the figures show that residential energy consumption and our different measures of climate are highly correlated. Figures 1.9 and 1.10 show the correlation of 5 and 10 years moving averages to residential energy consumption across different climatic regions and the correlations are quite promising. However, as we will see later, the number of observations for some regions if fairly low and that might be a possible issue in the estimation process. In a small section, later on, we will also use maximum and minimum temperatures to see if extreme temperatures have any effect on residential energy consumption.



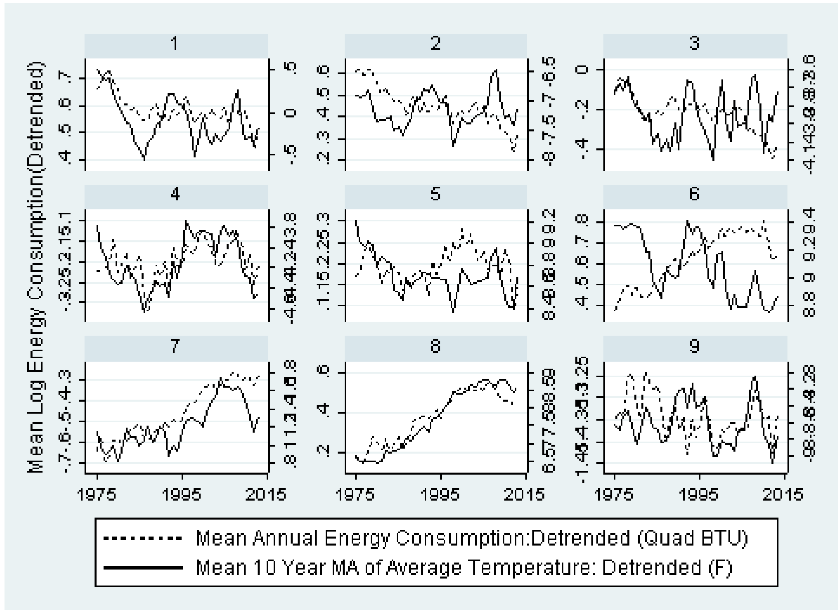
Notes: This figure represents the mean annual log residential energy consumption and the mean 5-year moving average of temperature, averaged across all states, for each year. The variables have been detrended in order to eliminate the time trend.

Figure 1.7: Relationship between Residential Energy Consumption and 5 Yr Moving Average Temperature



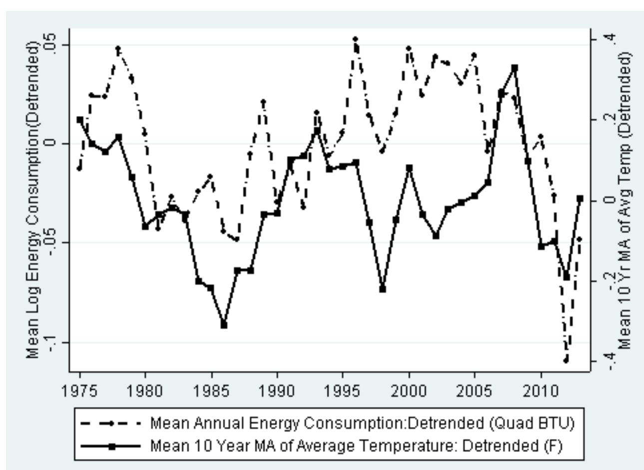
Notes: This figure represents the mean annual log residential energy consumption and the mean 5-year moving average of temperature, averaged across all states within a given climate region, for each year. The variables have been detrended in order to eliminate the time trend.

Figure 1.8: Relationship between Residential Energy Consumption and 5 Yr Moving Average Temperature by Climate Regions



Notes: This figure represents the mean annual log residential energy consumption and the mean 10-year moving average of temperature, averaged across all states within a given climate region, for each year. The variables have been detrended in order to eliminate the time trend.

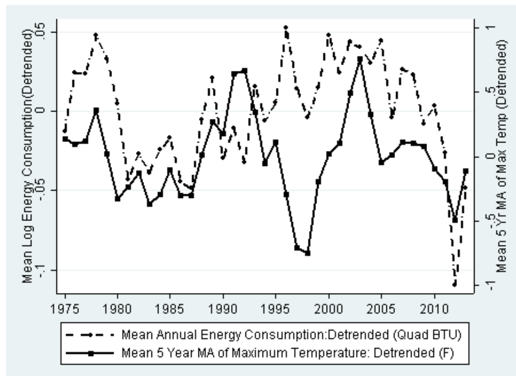
Figure 1.9: Relationship between Residential Energy Consumption and 10 Yr Moving Average Temperature by Climate Regions



Notes: This figure represents the mean annual log residential energy consumption and the mean 10-year moving average of temperature, averaged across all states, for each year. The variables have been detrended in order to eliminate the time trend.

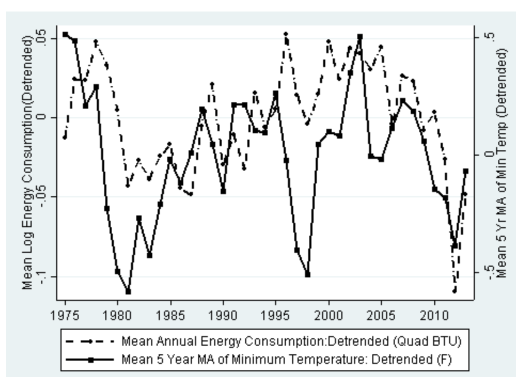
Figure 1.10: Relationship between Residential Energy Consumption and 10 Yr Moving Average Temperature

The following two figures (Figure 1.11 and Figure 1.12) shows the correlation of residential energy consumption with maximum and minimum temperature respectively. We will use the maximum and minimum temperature data in a minor analysis later.



Notes: This figure represents the mean annual log residential energy consumption and the mean 5-year moving average of daily maximum temperature, averaged across all states, for each year. The variables have been detrended in order to eliminate the time trend.

Figure 1.11: Relationship between Residential Energy Consumption and 5 Yr Moving Average of Maximum Temperature



Notes: This figure represents the mean annual log residential energy consumption and the mean 5-year moving average of daily minimum temperature, averaged across all states, for each year. The variables have been detrended in order to eliminate the time trend.

Figure 1.12: Relationship between Residential Energy Consumption and 5 Yr Moving Average of Minimum Temperature

1.5 Empirical Methodology

The conceptual framework of this paper is similar to Deschenes and Greenstone (2011) where they lay down a simple model to show how the effect of temperature on energy consumption enters the households willingness to pay/accept for a degree change in temperature. Since price data is available, one can multiple the estimated coefficient with the price to get the monetary value. This simple model does not completely capture what we are trying to do in this paper but serves a basic foundation from which we can build ours. The key idea of this paper is to segregate the effects of intensive margin adjustments and extensive margin adjustments in residential energy consumption in response to temperature changes. The objective for having bins of temperatures instead of the temperatures themselves is to see if there are any non-linear effects across different bins, that is to estimate if economic agents respond more in the intensive or the extensive margin due to extreme temperatures. We do this the extending Deschenes and Greenstone (2011) to incorporate climate trends as additional control in the model. This paper also deviates from the literature which uses "heating degree days (HDD) and "cooling degree days" (CDD) because this approach is not capable of capturing the non-linear impacts across different temperature bins.

In order to estimate the causal effect of temperature on residential energy consumption and get at our measure of adaptation to climate change, we exploit variation in contemporaneous daily average temperatures as well as the long term temperature trend (i.e. the climate normal), as given by the 5 year moving average of temperature. Following Deschenes and Greenstone (2011), and modifying their specification, we estimate the following econometric model:

$$\ln(C_{st}) = \sum_j \beta_j^{Cont} TMean_{stj}^{Cont} + \sum_l \gamma_l^{Prep} Prep_{stl} + \mathbb{X}_{st}\delta + \sum_j \beta_j^{MA} TMean_{stj}^{MA} + \alpha_s + \theta_{dt} + \epsilon_{st},$$

where s denotes state, t denotes year and d indexes the climate regions as defined by the National Oceanic and Atmospheric Administration (NOAA). The variable $TMean_{stj}^{Cont}$ denotes the *average number of days* in state s and year t, where the daily mean temperature is in the jth temperature bin. We have first calculated the number of days that daily mean temperature is in each of our 11 temperature bins, for each county-year and then averaged this to the state level. As discussed earlier, in order to control for historical temperature trends, we also control for $TMean_{stj}^{MA}$ which denotes the *average number of days* in state s and year t, where the *5 year moving average* of daily mean temperature is in the jth temperature bin. Hence, the only functional form restriction that we assume here is that the impact of current as well as moving average of temperature on residential energy consumption is constant within 10°F intervals. β_j^{Cont} captures the effect of an additional day that the *contemporaneous temperature* was in bin j whereas β_j^{MA} captures the effect of an additional day that the *historical average temperature* was in bin j. We also include state fixed effects (α_s), climate region by year fixed effects (θ_{dt}) and a stochastic error term (ϵ_{st}). In addition to this we also control for state population and GDP. [4]

The way we model temperature is as follows: *Contemporaneous temperature= Moving Average of Temperature (climate) +Shock (weather)*. In our specification, when one only controls for contemporaneous temperature, the coefficient captures two components. It captures the responses in energy consumed 1) if there is change is weather 2) if there is a change in the climate. The change in climate

[4]Additionally squared population and GDP have been controlled in the robustness section.

has two effects: it changes the contemporaneous temperature and it also induces changes made by the households in the extensive margin (for example, buying a new air conditioner). Therefore, in addition to contemporaneous temperature, when we control for the moving average of temperatures (climate), the coefficient of the contemporaneous temperature captures only response to a change in the shock (weather) component of the contemporaneous temperature. Hence, to get at the extensive margin adjustment we need to look at the difference in the estimates of the coefficients of the moving average (climate) and contemporaneous temperatures. β_j^{Cont} captures households reaction to contemporaneous change in temperatures and β_j^{MA} captures household's reaction to the trend of temperatures that they see in the previous years. So, the difference between the two estimates gives us a measure of adjustment in the extensive margin (adaptation to climate change). Suppose there is no adaptation to climate change in any way, then the effect of an additional day in some bin for the moving average and the contemporaneous weather should be the same and hence β_j^{Cont} should not be significantly different from β_j^{MA}. If, say for some reason, household before more environment conscious or are simply getting used to higher temperatures in summer or lower temperatures in winter, then we would expect $\beta_j^{Cont} > \beta_j^{MA}$. On the other hand, if people are just responding by buying more AC's or using more and more heating in winters, then we would expect $\beta_j^{Cont} < \beta_j^{MA}$.

1.6 Results

In this section, we will discuss the key findings of this paper. Firstly, we will look at pooled estimates for all the states in US for the years 1975-2013. These results will consider 5 years moving average as a measure of climate and we will also get into a detailed analysis of the regional impacts of temperature on energy consumption by households. Then we will move on to analyze how different notions of climate might affect our estimates. More specifically, if we change our notion of climate from 5 years moving average to 10 years moving average, then the key question is whether the results change and if so, then why do they change. This will be followed by a small section which will look at whether maximum temperatures or minimum temperatures have any substantial effect on the intensive or the extensive margin with regards to energy consumption.

1.6.1 Main Results

Table 1.3 shows the impact of changes in temperature on energy consumption by households for all the states in the US. The first two columns of the table do not include the 5 years moving average of temperatures and therefore looks at the effect of one more day in any of the 9 bins considered in the model on the energy consumption. Columns 3 and 4 shows results after controlling for the 5 years moving averages and as mentioned before it helps us in two ways. Firstly, controlling for the moving average or our notion of climate in the model, yields a better estimate of the contemporaneous effects and secondly it gives us scope of talking about how households react in the extensive margin. The coefficients reported in the table are relative to energy consumption on a day where the temperature is in the range of $50 - 60°F$ and we report the coefficients for the lowest two and the highest two bins of temperatures.

As it has been noted recently in the literature [5], we also find an inverted U shape in the response function of energy consumption to temperature. In

[5] Deschenes and Greenstone (2011) also find a non-linear response function.

Table 1.3: Main Estimates-Effect of Temperature on Residential Energy Consumption (5-Yr Moving Average)

Variables	(1)	(2)	(3)	(4)
No. of Days $TMean^{Cont} \leq 10°$ F	0.0044***	0.0047***	0.0043***	0.0045***
	(0.0009)	(0.0009)	(0.0007)	(0.0007)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0031***	0.0032***	0.0033***	0.0034***
	(0.0009)	(0.0008)	(0.0007)	(0.0007)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0023***	0.0023***	0.0019***	0.0019***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)
No. of Days $TMean^{Cont} > 90°F$	0.0043***	0.0042***	0.0032***	0.0031***
	(0.0014)	(0.0015)	(0.0010)	(0.0010)
No. of Days $TMean^{MA} \leq 10°F$			0.0039	0.0045*
			(0.0025)	(0.0024)
No. of Days $10°F < TMean^{MA} \leq 20°F$			0.0048**	0.0051***
			(0.0020)	(0.0019)
No. of Days $80°F < TMean^{MA} \leq 90°F$			0.0022*	0.0023*
			(0.0012)	(0.0012)
No. of Days $TMean^{MA} > 90°F$			0.0047**	0.0041*
			(0.0022)	(0.0023)
Population and GDP	Y	Y	Y	Y
Population x GDP	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

our preferred specifications (columns 2 and 4 respectively of table 1.3) we see that the response to changes in contemporaneous temperatures are higher for the extreme two bins of temperatures. We find that when we are not controlling for the climate (or moving averages) the coefficient of one extra day in any of the two extreme bins is in the range of $0.43 - 0.44$ and they are highly significant. As we have seen in the previous section, once we include the bins for moving averages, the coefficient of the contemporaneous temperatures will capture just the effect of the shock component in the temperature. Subtracting the coefficient of the contemporaneous temperatures from the coefficient obtained for the bins of the moving averages will give us a measure of extensive margin adaptation. Interestingly, when we control the moving average, the impact of contemporaneous temperatures on energy consumption for the highest bin drops to 0.0031 and noting that the coefficient of the moving average is 0.0041, we can say that when one does not control for the climate, a substantial portion of the effect of contemporaneous temperature on energy consumption is because of adjustments in the extensive margin. In our data, we find that the in response to change in climate, the coefficient which gives the impact on residential energy consumption 0.10 higher than it would be had there been no extensive margin adjustment for the highest temperature bin. We also find significant effects but of lower magnitude for the second highest temperature bin and the proportion of the effect on energy consumption which can be attributed to adjustments in the extensive margin is in similar range to what we found for the highest temperature bin. For the lowest two bins of temperatures, we find significant impacts of one more day in any of those bins and the coefficients of the moving averages suggest that there is no significant adjustment in the extensive margin for the lowest bin but there is significant adjustment for the bin with temperatures $10-20^{\circ}F$. This seems to suggest that households who experience one more day in the bin with temperatures less than $10^{\circ}F$ are already well equipped and need not make any changes in the extensive margin to combat lower temperatures. We will use the coefficients obtained for the contemporaneous temperatures after controlling for the moving averages to get an idea about how changes in temperatures will affect energy consumption at the end of the century.

1.6.2 Responses to Longer Term Climate Trends

In this section, we will expand our analysis to look at different notions of climate and how they influence our results. The primary motivation behind using the moving averages as additional controls is to isolate the adjustment to the extensive margin from the reaction in the intensive margin. One important aspect of our paper is the way we model contemporaneous temperature, we say that contemporaneous temperature is the moving average of temperature and some shock to the moving average. In this section, we want to see how agents react to longer term climatic trends since it is possible that the knowledge of a bigger past might influence agents long-run or extensive margin decisions. Table 1.4 shows the results when we consider 10 years moving average as a measure of climate. The first two columns give us the impacts of contemporaneous temperatures on energy consumption when we do not control for the moving averages. The last two columns give us the impact of one more day in a year in a particular temperature bin on energy consumption when one controls for the climate trends. In our preferred specification, which is column 4 of the table, we find that the coefficients for the contemporaneous temperature bins are highly significant and are in similar ranges to what we got when we used the 5 year moving averages. We also find significant effect for the moving average bin with temperatures between $10-20°F$ and we find bigger adjustments in the extensive margin in this case in comparison to the case when we controlled for 5 years moving average. This result is in line with the intuition that over a longer period agents would adjust more in the extensive margin. One should also note that we do not find significant results for the coefficients of the other moving average bins but still controlling for this has a substantial impact on the coefficients of the contemporaneous temperatures, as this helps in getting rid of bias one would have in the estimates of the coefficients of the contemporaneous temperature if one does take into account the extensive margin adjustments.

Table 1.4: Effect of Temperature on Residential Energy Consumption (10-Yr Moving Average)

Variables	(1)	(2)	(3)	(4)
No. of Days $TMean^{Cont} \leq 10°F$	0.0044***	0.0047***	0.0040***	0.0042***
	(0.0009)	(0.0009)	(0.0008)	(0.0008)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0031***	0.0032***	0.0030***	0.0031***
	(0.0009)	(0.0008)	(0.0007)	(0.0007)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0023***	0.0023***	0.0018***	0.0019***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)
No. of Days $TMean^{Cont} > 90°F$	0.0043***	0.0042***	0.0030***	0.0030***
	(0.0014)	(0.0015)	(0.0010)	(0.0010)
No. of Days $TMean^{MA} \leq 10°F$			0.0025	0.0035
			(0.0029)	(0.0029)
No. of Days $10°F < TMean^{MA} \leq 20°F$			0.0044*	0.0051**
			(0.0023)	(0.0022)
No. of Days $80°F < TMean^{MA} \leq 90°F$			0.0003	0.0004
			(0.0019)	(0.0018)
No. of Days $TMean^{MA} > 90°F$			0.0015	0.0008
			(0.0027)	(0.0027)
Population and GDP	Y	Y	Y	Y
Population x GDP	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

1.6.3 Climatic Regions

In this section, we have divided all the states into 9 different climatic regions to look at the heterogeneity of effects across different regions in the US and table 1.5 reports the estimates. We have considered the 5 years moving average as a measure of climate and report the results for our preferred specification where we also control population and GDP in addition to rainfall. The objective of this section is twofold. On one hand, it allows us to show that there is substantial variation in response of energy consumption to temperature changes and hence results from studies which focus on a particular state or region should be treated with caution as there might be issues of external validity. On the other hand, although limited to a certain degree due to low number of observations for each climatic region we are still able to get some interesting observations. We are not reporting results for the region *West* since the number of observations are too low and there is not much variation is the data for getting anything meaningful from the regressions. As noted by Deschenes and Greenstone (2011), for the regions, we find better results for the lower temperature bins in comparison to the higher temperature bins because of lack of observations in the higher temperature bins. For the bin with less than $10^{\circ}F$ we find significant effects in the coefficient for the contemporaneous temperatures in the regions Ohio Valley, Upper Midwest, North-East, South and Rockies with values ranging from 0.0019 – 0.0058. For the highest temperature bin (greater than $90^{\circ}F$) we have some evidence for the effect of contemporaneous temperatures in the Ohio Valley and Midwest. In terms of evidence for adjustments in the extensive margin, we find positive evidence in the North-East region. As mentioned before, the objective of this section was not to delve too much into the actual numbers because of data constraints. We think we have been able to show two things- firstly we show that there is variation across climatic regions with respect to reaction to temperature changes and also the lower values of coefficients found here in comparison to the literature shows controlling for the climate is important and helps us in getting better estimates of the contemporaneous effects.

Table 1.5: Estimates by Climate Region-Contemporaneous and 5 Yr Moving Average of Temperature

VARIABLES	Ohio Valley	Upper Midwest	Northeast	Northwest	South	Southeast	Southwest	Rockies
No. of Days $TMean^{Cont} \leq 10°F$	0.0032**	0.0019*	0.0040**	-0.0015	0.0058*	-0.0267**	0.0076	0.0044**
	(0.0011)	(0.0008)	(0.0018)	(0.0027)	(0.0023)	(0.0094)	(0.0046)	(0.0012)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0024***	0.0013	0.0054***	0.0037*	0.0040***	0.0081**	0.0025	0.0042**
	(0.0006)	(0.0011)	(0.0014)	(0.0012)	(0.0008)	(0.0030)	(0.0025)	(0.0010)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0012**	0.0001	0.0022	0.0024	0.0022*	0.0016*	0.0009	0.0022
	(0.0004)	(0.0007)	(0.0013)	(0.0047)	(0.0009)	(0.0006)	(0.0009)	(0.0016)
No. of Days $TMean^{Cont} > 90°F$	0.0108***	0.0836*	-0.0156	-0.0352	0.0008	0.0142	0.0011	-0.0822**
	(0.0025)	(0.0288)	(0.0209)	(0.0763)	(0.0013)	(0.0075)	(0.0029)	(0.0268)
No. of Days $TMean^{MA} \leq 10°F$	-0.0085*	-0.0028	0.0175***	0.0130	0.1013		-0.0080	0.0012
	(0.0041)	(0.0048)	(0.0050)	(0.0224)	(0.0530)		(0.0311)	(0.0045)
No. of Days $10°F < TMean^{MA} \leq 20°F$	-0.0032*	-0.0008	0.0144***	-0.0086*	0.0047	-0.0614	0.0036	0.0019
	(0.0013)	(0.0039)	(0.0028)	(0.0020)	(0.0048)	(0.0393)	(0.0038)	(0.0037)
No. of Days $80°F < TMean^{MA} \leq 90°F$	0.0005	-0.0004	0.0088**	0.0264**	0.0016	0.0017	0.0073*	-0.0076
	(0.0012)	(0.0119)	(0.0036)	(0.0032)	(0.0017)	(0.0017)	(0.0028)	(0.0085)
No. of Days $TMean^{MA} > 90°F$					-0.0099		0.0068**	
Observations	273	156	429	117	234	234	156	195

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Each climate region specification has been estimated separately. Since the West has only two states, and hence negligible variation, the coefficients have not been reported. For the Southeast, there was no variation in the bin for days below 10 F, and hence the coefficient has been dropped. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

1.6.4 Responses to Extreme Temperatures

In this section, we want to investigate whether economic agent's decision depends on its experience of extreme temperatures. The average temperature used in the previous section was computed using the daily maximum and minimum temperatures, so the average temperature will certainly incorporate the effect of a very high temperature or a very low temperature. But significant difference between the maximum and minimum temperatures might not fully capture if experiences of extremely temperatures has any effect on decision making of economic agents with regards to energy consumption.

Now the key question is the potential economic reason for such an effect. Extreme temperatures today might be an indicator for extreme temperatures tomorrow and households or economic agents might react in the extensive margin because of that. For example, in a certain place, in recent years the minimum temperature might be getting lower than what it usually is, then one might think of buying a new heater. Note that this affect would be captured to some extent when we use the average temperatures but instead if we use the maximum temperatures or minimum temperatures, we can get some idea whether apart from observing the average climate trends and making extensive margin adjustments, does extreme temperatures also play a role in adjustment in the extensive margin.

Therefore, in order to investigate whether extreme temperatures have any effect on residential energy consumption, we create bins for extreme temperatures. More specifically, using the same methodology, as explained in the data section for the average temperatures, we create bins of temperatures for the maximum temperatures and the minimum temperatures. For the maximum temperatures, we have bins from less than $50^{\circ}F$ to greater than $90^{\circ}F$ with intervals of $10^{\circ}F$ in between and for the minimum temperatures we have bins for less than $10^{\circ}F$ to greater than $50^{\circ}F$ with intervals of $10^{\circ}F$ in between. We use 5 years moving average as a notion of climate in this section. We separately run regressions with the maximum and minimum temperatures and we find significant impacts of extreme temperatures on residential energy consumption. Tables 1.6 and 1.7 report the estimated impacts of the maximum and minimum temperatures respectively on residential energy consumption. We have also constructed and controlled for bins for moving average of maximum and minimum temperatures. We find significant impacts of extreme temperatures on extensive margin adjustments made by households, which suggest that households not only react to average temperatures but extreme climate events can potentially play an important role in residential energy consumption. It is worth noting that we should not be expecting the coefficients of the contemporaneous maximum and minimum temperatures to be similar to the coefficients of the contemporaneous average temperatures since the maximum and minimum only captures a part of the average temperatures. Additionaly, we also find some evidence of extensive margin adjustment for maximum temperatures in the Northeast and Southwest regions. For the minimum temperature, we find some similar impact for the Northeast region. Tables 1.8 and 1.9 report the results for the regional impacts for maximum and minimum temperatures respectively.

Table 1.6: Effect of Maximum Temperature on Residential Energy Consumption

VARIABLES	(1)	(2)	(3)	(4)
No. of Days $TMax^{Cont} \leq 50°F$	0.0021***	0.0021***	0.0022***	0.0022***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)
No. of Days $TMax^{Cont} > 90°F$	0.0012***	0.0012***	0.0011***	0.0011***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
No. of Days $Tax^{MA} \leq 50°F$			0.0028**	0.0027**
			(0.0012)	(0.0012)
No. of Days $TMax^{MA} > 90°F$			0.0020**	0.0020**
			(0.0008)	(0.0008)
Population and GDP	Y	Y	Y	Y
Population x GDP	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Regressions report responses to daily maximum temperatures and 5 year moving averages of the same. $10°F$ bins have been constructed based on the distribution of daily maximum temperatures. Estimates of the two most extreme bins have been reported. All estimates are relative to the intermediate bin of 70-80$°F$. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

Table 1.7: Effect of Minimum Temperature on Residential Energy Consumption

VARIABLES	(1)	(2)	(3)	(4)
No. of Days $TMin^{Cont} \leq 10°F$	0.0024***	0.0025***	0.0025***	0.0026***
	(0.0007)	(0.0007)	(0.0005)	(0.0005)
No. of Days $TMin^{Cont} > 60°F$	0.0013**	0.0012**	0.0008**	0.0008**
	(0.0005)	(0.0005)	(0.0004)	(0.0004)
No. of Days $TMin^{MA} \leq 10°F$			0.0032**	0.0036**
			(0.0016)	(0.0015)
No. of Days $TMin^{MA} > 60°F$			0.0028***	0.0027**
			(0.0010)	(0.0010)
Population and GDP	Y	Y	Y	Y
Population x GDP	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Regressions report responses to daily minimum temperatures and 5 year moving averages of the same. $10°F$ bins have been constructed based on the distribution of daily minimum temperatures. Estimates of the two most extreme bins have been reported. All estimates are relative to the intermediate bin of 30-40$°F$. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

Table 1.8: Estimates by Climate Region-Contemporaneous and 5 Yr Moving Average of Maximum Temperature

VARIABLES	Ohio Valley	Upper Midwest	Northeast	Northwest	South	Southeast	Southwest	Rockies
No. of Days $TMax^{Cont} \leq 50°F$	0.0009	0.0013	0.0044**	-0.0005	0.0035**	0.0013**	0.0023**	0.0013
	(0.0009)	(0.0009)	(0.0016)	(0.0008)	(0.0012)	(0.0003)	(0.0006)	(0.0007)
No. of Days $TMax^{Cont} > 90°F$	0.0006**	0.0009	0.0017**	0.0015	0.0025**	0.0020***	0.0006	-0.0003
	(0.0002)	(0.0010)	(0.0007)	(0.0015)	(0.0007)	(0.0003)	(0.0003)	(0.0011)
No. of Days $TMax^{MA} \leq 50°F$	0.0016	0.0021	0.0045	0.0014	-0.0023	-0.0005	0.0024	0.0017
	(0.0016)	(0.0033)	(0.0035)	(0.0012)	(0.0025)	(0.0013)	(0.0030)	(0.0050)
No. of Days $TMax^{MA} > 90°F$	-0.0003	0.0004	0.0065**	0.0097	0.0003	0.0017	0.0057***	-0.0011
	(0.0011)	(0.0085)	(0.0026)	(0.0036)	(0.0022)	(0.0014)	(0.0009)	(0.0029)
Observations	273	156	429	117	234	234	156	195

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Each climate region specification has been estimated separately. $10°F$ bins have been constructed based on the distribution of daily maximum temperatures. Estimates of the two most extreme bins have been reported. All estimates are relative to the intermediate bin of 70-80$°F$. Since the West has only two states, and hence negligible variation, the coefficients have not been reported. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

Table 1.9: Estimates by Climate Region-Contemporaneous and 5 Yr Moving Average of Minimum Temperature

VARIABLES	Ohio Valley	Upper Midwest	Northeast	Northwest	South	Southeast	Southwest	Rockies
No. of Days $TMin^{Cont} \leq 10°F$	0.0020	0.0020	0.0023	0.0005	0.0032*	0.0059	0.0027	0.0027**
	(0.0010)	(0.0013)	(0.0018)	(0.0005)	(0.0013)	(0.0030)	(0.0014)	(0.0008)
No. of Days $TMin^{Cont} > 60°F$	0.0001	-0.0002*	0.0002	0.0042**	-0.0004	-0.0001	-0.0014	0.0006
	(0.0007)	(0.0001)	(0.0008)	(0.0005)	(0.0009)	(0.0005)	(0.0008)	(0.0008)
No. of Days $TMin^{MA} \leq 10°F$	-0.0031	0.0008	0.0114***	-0.0083*	0.0072	-0.1659**	0.0074	-0.0006
	(0.0017)	(0.0015)	(0.0015)	(0.0024)	(0.0054)	(0.0560)	(0.0046)	(0.0015)
No. of Days $TMin^{MA} > 60°F$	0.0017**	0.0025	0.0019	0.0191***	0.0021	0.0029	0.0013	0.0008
	(0.0006)	(0.0029)	(0.0021)	(0.0018)	(0.0035)	(0.0015)	(0.0017)	(0.0025)
Observations	273	156	429	117	234	234	156	195

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Each climate region specification has been estimated separately. $10°F$ bins have been constructed based on the distribution of daily minimum temperatures. Estimates of the two most extreme bins have been reported. All estimates are relative to the intermediate bin of 30-40$°F$. Since the West has only two states, and hence negligible variation, the coefficients have not been reported. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

1.7 Robustness

In this section, we will modify our primary specification in a number of different ways and see if our results are robust to minor tweaks in the model. Broadly, we will make changes to the specification in three different dimensions. As has been noted in the data section, we considered weather stations within a distance of 200 kilometers from the county centroid for constructing the weather variables. We will consider three variations with respect to the distance unto which the weather stations will be considered for construction of the climate variables. Next, we will consider a different notion of climate. In our current specification, we assume that households get one year to make extensive margin adjustments. We would change this model and model climate such that they have more time to make adjustments to the extensive margin and we will do this by controlling for lagged moving averages. Finally, we will further make changes to our specification by including more controls which shall capture the non-linear impacts of population and *GDP* and their interaction on residential energy consumption.

The main results reported in the previous section use weather variables which were constructed by considering weather stations which are within 200 kilometers from the centroid of a country. After that bins for the weather variables were created and aggregated to the state level by using county population as weights. We change the radius and consider weather stations which are within 100, 150 and 250 kilometers respectively from the centroid of the county. A similar procedure has been followed to aggregate the data at the state level. Table 1.10 reports the estimates for the three different radii, columns 1, 3 and 5 report estimates for the three cases when we do not control for the moving averages and columns 2, 4 and 6 report estimates for the three cases when we control for the moving averages. As we did for our preferred specification in the previous section, we consider 5 years moving average as a measure of climate. As we see in table three our estimates are robust to different variations in the construction of the weather variables. Estimates for both the contemporaneous temperatures and for the moving averages do not vary much across different radii and they are not significantly different from the estimates we had for our preferred specification (see table 1.3) in the main results. As we had seen in the results section, there is significant impacts for all the four bins reported in table 1.10, for both the contemporaneous temperatures and the moving averages. For the lower end of the temperature distribution, we find evidence of adjustment in the extensive margin because of one more day in the bin with temperatures between $10 - 20°F$. For the other extreme temperature bins, we find evidence of extensive margin adjustment for both the bins reported. The change is radii changes the number of weather stations considered in the sample. Overall, we find that our results do not change significantly when we reduce or increase the number of weather stations considered in our specification.

Table 1.10: Robustness- Alternative Criteria for Selection of Weather Stations (5 Yr Moving Average)

	100 km radius		150 km radius		250 km radius	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
No. of Days $TMean^{Cont} \leq 10°F$	0.0046***	0.0045***	0.0046***	0.0044***	0.0047***	0.0045***
	(0.0009)	(0.0007)	(0.0009)	(0.0007)	(0.0009)	(0.0007)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0031***	0.0033***	0.0031***	0.0033***	0.0033***	0.0035***
	(0.0008)	(0.0007)	(0.0008)	(0.0007)	(0.0008)	(0.0007)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0023***	0.0019***	0.0023***	0.0019***	0.0023***	0.0019***
	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0004)
No. of Days $TMean^{Cont} > 90°F$	0.0042***	0.0028***	0.0042***	0.0030***	0.0041***	0.0031***
	(0.0015)	(0.0010)	(0.0015)	(0.0010)	(0.0015)	(0.0011)
No. of Days $TMean^{MA} \leq 10°F$		0.0046*		0.0044*		0.0045*
		(0.0024)		(0.0024)		(0.0024)
No. of Days $10°F \leq TMean^{MA} < 20°F$		0.0053***		0.0051***		0.0052***
		(0.0019)		(0.0019)		(0.0019)
No. of Days $80°F < TMean^{MA} \leq 90°F$		0.0023*		0.0022*		0.0022*
		(0.0011)		(0.0011)		(0.0012)
No. of Days $TMean^{MA} > 90°F$		0.0041*		0.0040*		0.0042*
		(0.0023)		(0.0023)		(0.0024)
Population and GDP	Y	Y	Y	Y	Y	Y
Population x GDP	N	Y	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Columns (1), (3) and (5) correspond to Column (2) of the main results whereas Columns (2), (4) and (6) correspond to Column (4) of the main results. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

Next, we change the time a household has to make adjustments to the extensive margin and see how the results change. The way we have modeled climate in this paper is that we have controlled for 5 year moving averages in addition to the contemporaneous temperatures. Previously, we were controlling for contemporaneous temperatures of year t (say) and additionally we were controlling for 5 years moving average of temperatures which is for the period $t-1$ to $t-5$. Now, we will consider 5 years moving averages but will compute that using temperatures for periods $t-4$ to $t-8$, which implies that to control for 3 years lagged 5 years moving average. One implication this change has on the model is the fact that it changes the time the households get to adjust in the extensive margin. Previously, our model allowed for changes in climate to affect residential energy consumption through a one year lag. On the contrary, this model allows households to react to changes in climate through a three year lag. Intuitively, it is not straight forward to predict the behavior of households in this scenario. On the one hand, it might be the case that households get more time to react when there is a two year lag and hence one should expect more adjustment in the extensive margin since households have more time to adjust. On the other hand, lagged moving average might not be a good representation for climate in our model because it might be more difficult for households to think about lagged moving averages and hence we might expect a lower extensive margin adjustment when we control for lagged moving averages in comparison to 1 year lagged moving averages. Another thing that should be noted here is that the estimates for the contemporaneous temperature bins when we control for lagged moving averages should not be significantly different to the scenario where we do not have the lag.

Table 1.11 reports the results where we consider three different lags in the model, more specifically 2 years, 3 years and 4 years lagged moving average. The results show that as one would expect, the estimates of the impact of one more day in the extreme temperature bins of contemporaneous temperatures are not significantly different from the estimates in table 1.3 where we had a one year lag. Columns 2, 4 and 6 table 1.11 report estimates for 2, 3 and 4 years lagged moving average for our preferred specification. As for the estimates of the moving average bins, we do not find any significant effect expect for the bin of $10 - 20°F$. This shows that lagged moving average might not be a good representation of climate and the fact that the difference between the coefficient of moving average and contemporaneous temperatures decreases in comparison to the difference obtained in column 4 of table 1.3 shows that households might not remember lagged moving average that well and hence it might not capture climate properly. Moreover, this is further supported by the fact that the coefficient of the moving average bin of $10 - 20°F$ decreases when we increase the lag.

Finally, we include squared values of population and *GDP* in our specification in order to check if there are non-linear effects of population and GDP on residential energy consumption and if that is confounding our results in anyway. Table 1.12 reports the results after adding these two additional controls. Column 3 in table 1.12 shows that results do not change significantly in comparison to our preferred specification in table 1.3 except for the moving average of the bin with temperatures greater than $90°F$. Column 4 of table 1.12 shows that if we also control for the interaction of squared population and GDP, then we find significant impact for the contemporaneous temperatures and for the bins $10 - 20°F$ and $20 - 90°F$. This suggests that our main results in table 1.3 might have overestimated the effect of climate on residential energy consumption but nonetheless significant impacts in the other moving average bins shows that it is important to control for climate in the specification. This might also suggest that households are already well-equipped to cope with an additional day in the two extreme temperature bins and hence they do not need to make any adjustments in the extensive margin. This shows that even after controlling for squared population and *GDP* and their interaction, we still find evidence of extensive margin adjustment and our results for the contemporaneous temperatures are quite stable across specifications.

Table 1.11: Robustness- Lagged Responses to Climate (5 Yr Moving Average)

VARIABLES	2 Year Lag		3 Year Lag		4 Year Lag	
	(1)	(2)	(3)	(4)	(5)	(6)
No. of Days $TMean^{Cont} \leq 10°F$	0.0044***	0.0046***	0.0044***	0.0046***	0.0043***	0.0045***
	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0031***	0.0032***	0.0029***	0.0030***	0.0030***	0.0031***
	(0.0007)	(0.0007)	(0.0008)	(0.0007)	(0.0008)	(0.0008)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0019***	0.0019***	0.0019***	0.0020***	0.0020***	0.0021***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
No. of Days $TMean^{Cont} > 90°F$	0.0030***	0.0029***	0.0032***	0.0032***	0.0035***	0.0034***
	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0010)	(0.0010)
No. of Days $TMean^{MA} \leq 10°F$	0.0028	0.0034	0.0013	0.0020	0.0003	0.0010
	(0.0025)	(0.0025)	(0.0023)	(0.0022)	(0.0020)	(0.0020)
No. of Days $10°F \leq TMean^{MA} < 20°F$	0.0041**	0.0045**	0.0036*	0.0040**	0.0030*	0.0035**
	(0.0019)	(0.0019)	(0.0018)	(0.0017)	(0.0016)	(0.0016)
No. of Days $80°F < TMean^{MA} \leq 90°F$	0.0019	0.0020*	0.0014	0.0015	0.0009	0.0010
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0013)	(0.0014)
No. of Days $TMean^{MA} > 90°F$	0.0042*	0.0036	0.0029	0.0023	0.0015	0.0010
	(0.0022)	(0.0023)	(0.0024)	(0.0025)	(0.0026)	(0.0025)
Population and GDP	Y	Y	Y	Y	Y	Y
Population x GDP	N	Y	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872	1,872	1,872

Notes: Regressions include state fixed effects and climate region-by-year fixed effects. Results report responses to 5 year moving averages, lagged by 2 years, 3 years and 4 years respectively. Columns (1), (3) and (5) correspond to Column (3) of the main results whereas Columns (2), (4) and (6) correspond to Column (4) of the main results. Standard errors are clustered at the state level. ***, ** and * represent statistical significance at the 1%, 5% and 10% level respectively.

Table 1.12: Robustness- Quadratic controls for Population and GDP (5 Yr Moving Average)

VARIABLES	(1)	(2)	(3)	(4)
No. of Days $TMean^{Cont} \leq 10°F$	0.0047***	0.0044***	0.0046***	0.0043***
	(0.0008)	(0.0008)	(0.0007)	(0.0006)
No. of Days $10°F \leq TMean^{Cont} < 20°F$	0.0033***	0.0033***	0.0035***	0.0035***
	(0.0008)	(0.0007)	(0.0007)	(0.0006)
No. of Days $80°F < TMean^{Cont} \leq 90°F$	0.0023***	0.0024***	0.0019***	0.0020***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)
No. of Days $TMean^{Cont} > 90°F$	0.0040**	0.0033**	0.0029**	0.0025**
	(0.0016)	(0.0015)	(0.0011)	(0.0011)
No. of Days $TMean^{MA} \leq 10°F$			0.0046*	0.0039
			(0.0024)	(0.0024)
No. of Days $10°F \leq TMean^{MA} < 20°F$			0.0055***	0.0049**
			(0.0020)	(0.0019)
No. of Days $80°F < TMean^{MA} \leq 90°F$			0.0023*	0.0025**
			(0.0012)	(0.0011)
No. of Days $TMean^{MA} > 90°F$			0.0034	0.0016
			(0.0027)	(0.0029)
Population and GDP	Y	Y	Y	Y
Population x GDP	N	Y	N	Y
Population Sq and GDP Sq	Y	Y	Y	Y
Pop Sq x GDP Sq	N	Y	N	Y
Observations	1,872	1,872	1,872	1,872

1.8 Conclusion

In this paper, we have used a novel methodology to look at the intensive and extensive margin adjustments made by households or economic agents to cope with changes in temperature. Our approach allows us to get better estimates for the intensive margin adjustment because we are explicitly controlling for the climatic trends. For our preferred specification, our estimates yield higher coefficients for contemporaneous temperature for the lowest two bins in comparison to Deschenes and Greenstone (2011), who use a similar specification but do not control for climatic trends. Also, as we see in the results section, there are significant differences in the coefficients of the contemporaneous temperature when one controls for the climate trend in comparison to the case where one does not control for it. This clearly shows that it is important to control for the climatic trends and that would give us better estimates of change in residential energy consumption due to change in weather (shock to climate trends) and also allow us to look at adjustments in the extensive margin by looking at the difference between the coefficients of the moving average (climate trends) and corresponding coefficient of contemporaneous temperature bin. We also document a non-linear relationship between temperature changes and residential energy consumption. The various climate models predict that by the end of the century, due to global warming, we will have higher number of days in the higher temperature bins and lower number of days in the lower temperature bins. If one uses our estimates in calculating the monetary costs of climate change, then given that our estimates for the lower temperature bins are greater than Deschenes and Greenstone (2011), one would possibly find a lower monetary cost than their estimates.

We also use 10 years moving average instead of 5 years to see how the estimated impact changes with different notions of climate. We find that the effect on residential energy consumption due to a change in weather does not change in comparison to the case when we used 5 years moving average as a measure of climate. On the adjustment in the extensive margin we find significant impact for the bin with temperatures between $10-20°F$ and as expected, because of a bigger climate trend we find a bigger adjustment in the extensive margin.

We also find significant impacts for multiple climatic regions for both the lower and the second highest bin. There is also significant variation in impacts across climatic regions which suggests that results based on some region might not be extendable to other regions. We should also note that since we are using state year level data, the regressions involving different climatic regions lose number of observations and also lose variation in data in the highest bin.

Lastly, we look at the possibility of extreme temperatures affecting extensive margin adjustment. Instead of using bins of average temperatures, we construct bins of maximum and minimum temperatures and run separate regressions for the maximum temperature and minimum temperature bins. We have some preliminary evidence which suggests that extreme temperatures do affect residential energy consumption in the extensive margin. This suggests that households not only react to average temperatures but more extreme temperature realizations in the past might induce households to form an idea of the future average temperatures, which will in turn affect their extensive margin adjustments.

Overall, this paper helps to address an important issue in the literature. The reduced form estimates with contemporaneous temperature as an explanatory variable is not able to isolate extensive margin adjustments from intensive margin adjustments. Our methodology contributes to the literature in filling the void by getting at exact estimates for responses to changes in weather and changes due to extensive margin adjustment.

1.9 References

1. Akpinar-Ferrand, E., and Singh, A. (2010). “Modeling Increased Demand of Energy for Air Conditioners and Consequent CO2 Emissions to Minimize Health Risks Due to Climate Change in India”, *Journal of Environmental Science and Policy*, Volume 13, pp. 702-712.

2. Auffhammer, M., and Aroonruengsawat, A. (2011). “Simulating the Impacts of Climate Change, Prices and Population on California’s Residential Electricity Consumption”, Climatic Change, Volume 109, pp. 191-210.

3. Auffhammer, M., and Aroonruengsawat, A. (2012b). Erratum to: Simulating the Impacts of Climate Change, Prices and Population on California’s Residential Electricity Consumption”, *Climatic Change*, Volume 113, pp. 1101-1104.

4. Crowley, C., and Joutz, F. L. (2003). “Hourly electricity loads: Temperature elasticities and climate change”, *In 23rd U.S. Association of Energy Economics North American Conference*.

5. Dell, M., Jones, B. F., and Olken, B. A. (2009). “Temperature and Income: Reconciling New Cross-Sectional and Panel Estimates”, *American Economic Review: Papers and Proceedings* Volume 99, pp. 198-204.

6. Dell, M., Jones, B. F., and Olken, B. A. (2012). “Temperature Shocks and Economic Growth: Evidence from the Last Half Century”, *American Economic Journal: Macroeconomics* Volume 4, pp. 66-95.

7. Dell, M., Jones, B. F., and Olken, B. A. (2014). “What Do We Learn from the

Weather? The New Climate-Economy Literature", *Journal of Economic Literature* Volume 52, pp. 740-798.

8. Deschenes, O., and Greenstone, M. (2011). "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US", *American Economic Journal: Applied Economics*, Volume 3, pp. 152-185.

9. Energy Information Administration. (2011). "Air Conditioning in Nearly 100 Million U.S. Homes", see *http://www.eia.gov/consumption/residential/reports/air conditioning09.cfm*

10. Eskeland, G. S., and Mideksa, T. K. (2010). "Electricity Demand in a Changing Climate", *Mitig Adapt Strateg Glob Change*, Volume 15, pp. 877-897.

11. Franco, G. and Sanstad, A., (2008). "Climate Change and Electricity Demand in California", *Climatic Change*, Volume 87, pp. 139-151.

12. Mansur, E. T., Mendelsohn, R., and Morrison, W. (2008). "Climate Change Adaptation: A Study of Fuel choice and Consumption in the US Energy Sector", *Journal of Environmental Economics and Management*, Volume 55, pp. 175-193.

13. Rapson, D. (2014). "Durable Goods and Long-Run Electricity Demand: Evidence from Air Conditioner Purchase Behavior", *Journal of Environmental Economics and Management*, Volume 68, pp. 141-160.

14. Solon, G. (1992). "Intergenerational Income Mobility in the United States", *American Economic Review* Volume 82, pp. 393-408.

15. Vaage, K. (2000). "Heating Technology and Energy Use: a Dis-

crete/Continuous Choice Approach to Norwegian Household Energy Demand", *Energy Economics*, Volume 22, pp. 649- 666.

CHAPTER 2

EFFECT OF RISING INCOME INEQUALITY ON THE EQUILIBRIUM OF A CITY

2.1 Introduction

Rising economic inequality has been a major concern of policy makers since several countries around the globe are experiencing a rise in wealth as well as income inequality. The key concern is whether the rising differences affect different sections of the society disproportionately because of differences in the level of vulnerability between the rich and the poor. One of the key areas is the housing market, and the common understanding is that since land is scarce, economic inequality would lead to the situation where the rich would outbid the poor even more than before and therefore reduce the welfare of the poor.

One major issue in the developing countries is the large section of squatter population or informal settlements in the cities. IPCC report 2014 clearly states the magnitude of the problem as there has been a rapid increase in the number of new cities recently. Poor people move to the cities for better employment opportunities but have to live in very small poorly constructed houses. One might think that the papers in the literature which try to capture homelessness etc. captures informal settlements in the sense that they are also poor, but I believe there is something more to the problem of informal settlements than the fact they are also poor.

As the connotation informal settlements suggests that there are basic differences in the nature of their housing. Although there are several differences in the way we understand and treat informality, one way to understand informality is that informal housing does not come under the housing regulations and hence are of extremely bad quality and lack all basic amenities. Also, in many cities in India, informal settlements grow on free government land or unclaimed lands, so sometimes there is no direct competition over a piece of land between the rich and the poor. Many of the papers in the literature assume that both rich and poor are competing in the same market and that may be true in developed countries where people cannot live illegally but that is not the case in developing or underdeveloped countries because poor people do live illegally in very bad conditions.

The governments in various countries has tried out and continue to formulate policies aimed at decreasing slum population. Moreover, it is not even clear about which slum redevelopment policy works because it keeps on changing both with changes in government and within the regime of a government. It is well documented that people move to cities hoping for better opportunities but the number of people migrating is too large compared to how much the city can accommodate. Therefore poor people are forced to live in the slums. Even if government policies aimed at reducing slum population has some positive effect, the problem intensifies because the number of cities is growing at a rapid rate over the years and the IPCC projections of number of cities are also alarming. Figure 2.1 and 2.2 shows number of cities in 1950 and projection for 2025 and the difference is huge.

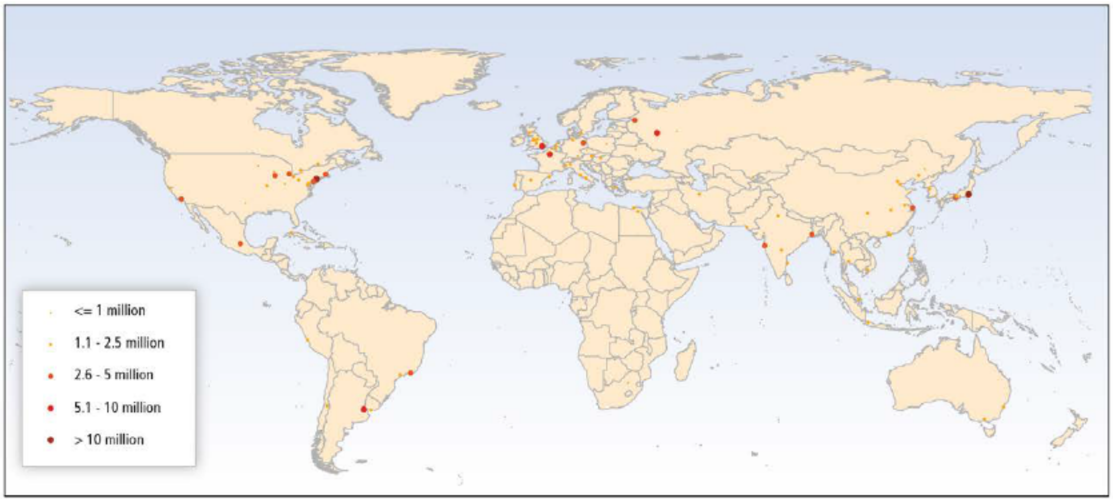


Figure 2.1: Cities in 1950: Source IPCC Report 2014

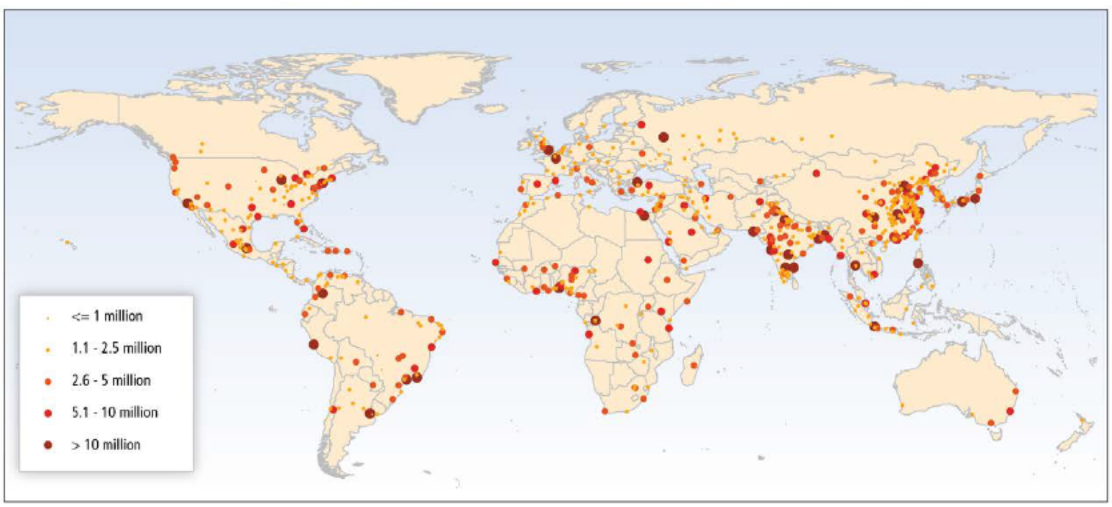


Figure 2.2: Projected Cities in 2025: Source IPCC Report 2014

According to UN-HABITAT, around 33% of the urban population in the developing world in 2012, or about 863 million people, lived in slums. Census data projections show that India's slum population will surge to 104 million by 2017 - or around 9% of the total projected national population of 1.28 billion that year. Census data also shows how a large section of the urban population lives in slums in the four metropolitan cities as high as 41.3% in Greater Mumbai, 29.6% in Kolkata, 28% in Chennai and about 15% in Delhi.

So I think, the rapid growth of cities, along with the rise in income inequality, has the potential to create a big problem in the developing countries and it is important to understand how the rise in income inequality affects the housing market equilibrium in the presence of slum population. Policies aimed at just reducing slum population might not work if it is the case that all the benefits of the policy are outdone by the changes in income inequality.

Therefore, I think it is important to understand that the presence of informal settlements is a slightly different problem than just having a poor population in the city. With rising income inequality, the demand for land from the rich might go up which can potentially lead to an increase in the price of existing houses and also a push for housing developers to build new houses. This would, in turn, lead to less land for the informal settlers and they would be forced to relocate to other places or live at even smaller houses. Another possible effect might be the poor who previously chose to live in formal housing, might relocate to slums if the price of housing in the formal sector goes up. So, in this paper, I want to set-up a theoretical framework, through which we can potentially explain and understand the above mentioned issues in a more structured way.

2.2 Literature Review

There are both theoretical and empirical studies which have looked at the issue of income inequality and its effects on housing markets. Mattanen and Tervio (2014) considers an assignment model where households vary by incomes and houses by the quality and shows that the impact of increased income inequality on house prices depends on the shapes of the distributions. Moreover, they find that during 1998-2007, higher income inequality led to lower average housing price in major metropolitan areas in USA. Matlack and Vigdor (2006) looks at metropolitan areas in USA from 1970-2000 and they find that when vacancies are low, increase in income of the rich lead to higher rents per room. Landvoigt, Piazzesi and Schneider (2015) show that in San Diego county, the major cause of a boom in housing prices in the 2000's was because of cheap credit available to the poor.

Another class of models widely used in the urban literature is the monocentric city models which were developed in the 1960's and 1970's (see Alonso (1964), Mills (1967) and Muth (1969)) but the main features of these models are the central business districts where people come to work and therefore living further away from the centre is costly because of the presence of construction costs. Another class of models introduced by Brueckner and Selod (2009), where they explicitly model formal and informal residents and look at the effects of formalization policies on the welfare of the rich and poor.

There have been work on homelessness in USA as well. Quigley, Raphael, and Smolensky (2001) show that variation in homelessness in US metropolitan cities can be explained to a large extent by increased demand for low quality housing and prices of housing. In another study by Mansur, Quigley, Raphael, and Smolensky (2002) show well known housing subsidy policies can help in eliminating homelessness in USA.

Glazer etal. (2008) tries to understand how income redistribution can lead to migration of people across counties and how it changes property values of the counties. Nieuwerburgh and Weill (2010) builds a general equilibrium model to explain why for a period of 30 years there has been an increase in the price of housing and also an increase in the variance of housing prices in US metropolitan areas. Gyourko etal. (2013) shows that variation and appreciation of housing prices in US metropolitan areas is a result of inelastic land supply and increase in rich households at the national level.

2.3 Model

We consider a simple model taken from Brueckner and Selod (2009) and then extend it later on to incorporate other things we need in this paper. Even in the basic model, we extend Brueckner's set-up of two income groups to three because that allows us to talk about the increase in income inequality between different groups and lets us see the effect of it not only on the two groups involved but on the other group as well. Let the size of land be denoted by $\bar{L}$ and the total land is divided into two parts L_s and L_f, where L_s and L_f denote the land occupied by the squatter population and land occupied by formal residents respectively. The population size is denoted by N and is divided into three parts N_s, N_{f_1} and N_{f_2} such that

$$N_s + N_{f_1} + N_{f_2} = N$$

where N_s is the number of individuals living in the informal or squatter land and N_{f_1} and N_{f_2} are the number of individuals living in the formal part of the land. It is important here to distinguish between formal and informal land. In the formal land the residents have to pay rent and in the informal land residents do not pay rent but have to pay an amount to the squatter organizer. N_{f_1} and N_{f_2} distinguishes two groups of formal residents with different income levels.

Each individual has some income which is exogenous to our model. The squatter population has income I_3, N_{f_1} individuals have income I_1 and N_{f_2} individuals have income I_2 such that

$$I_1 > I_2 > I_3$$

Individuals consume two goods, housing (q) and non-housing good (x) and have the following utility function

$$x^{1-\alpha} q^{\alpha}$$

There is cost of eviction which is given by

$$e(A, N_s, k) = AN_s k$$

where A denotes the payment each squatter households makes to the slum organizer and k captures the political difficulties of eviction, with both k and A being positive.

The number of people willing to stay in the formal sector is fixed and hence denoted by $\bar{N}_{f_1}$ and $\bar{N}_{f_2}$. Let p_f denote the price or rent of housing in the formal sector. Therefore the total demand for housing in the formal sector is the following

$$d_f(p_f) = \bar{N}_{f_1}[\frac{\alpha I_1}{p_f}] + \bar{N}_{f_2}[\frac{\alpha I_2}{p_f}]$$

The squatter organizer is assumed to of benevolent nature and hence tries to maximize the utility of the informal residents. The organizer takes into account the demand for land by the formal residents and assures no eviction in the equilibrium. The squatter population budget constraint is

$$A + x = I_3$$

where the price of the non-housing good has been normalized to 1 and therefore the problem faced by the organizer is as follows

$$\begin{aligned} \max{}_{\{A,q_s,N_s\}}(I_3 - A)^{1-\alpha} q_s^\alpha \\ \text{subject to } L_s + L_f = \bar{L} \\ N_s q_s = L_s \\ \bar{N}_{f_1}[\frac{\alpha I_1}{p_f}] + \bar{N}_{f_2}[\frac{\alpha I_2}{p_f}] = L_f \\ AN_s k \geq p_f \end{aligned}$$

The last constraint will hold with equality in equilibrium. Moreover, using this to replace p_f in the third constraint and then solving for L_s from the first constraint we can use the second constraint to compute q_s. Therefore the problem can be rewritten as (planner chooses N_s and A)

$$\max\ (I_3 - A)^{1-\alpha}(\frac{\bar{L} - \beta/AN_s}{N_s})^\alpha \tag{2.1}$$

where $\beta = \frac{(\bar{N}_{f_1}\alpha I_1 + \bar{N}_{f_2}\alpha I_2)}{k}$.

Solving the problem would yield the following solution

$$\begin{aligned} p_f^* &= \frac{2\alpha(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)}{\bar{L}} \\ A^* &= \alpha I_3 \\ q_s^* &= \frac{k I_3 \bar{L}^2}{4(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)} \\ N_s^* &= \frac{2(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)}{k I_3 \bar{L}} \\ L_s^* &= L_f^* = \bar{L}/2 \end{aligned}$$

Moreover, for this equilibrium to be sustained the parameters of the model should be such that the indirect utility of a squatter household from moving into the formal residential area should be less than the indirect utility received when living on the informal land. This condition implies

$$[(1-\alpha)y_3]^{1-\alpha}[\alpha y_3/p_f^*]^{\alpha} < [\frac{(y_3 - A)^{1-\alpha} k y_3 \bar{L}^2}{4(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)}]^{\alpha}$$

which implies $k\bar{L} > 2$.

2.3.1 Effects of Rising Income Inequality

In this section, we want to discuss how the housing market equilibrium is affected by changes in income inequality in two ways. First, we will ask what happens if income inequality widens between residents in the formal sector only. Second, we will ask what happens if income inequality widens between residents in the formal and the informal sector. We characterize widening of income inequality by looking at the ratios of income of two different income groups.

Case 1: Suppose I_1/I_2 increases, that is rich people in the formal sector become richer and also we consider and mean preserving redistribution, which implies

$$\frac{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2) + 2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)/k\bar{L}}{\bar{N}_{f_1} + \bar{N}_{f_2} + 2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)/kI_3\bar{L}}$$

does not change. Therefore, note that I_1 increases and I_2 decreases but since this is a mean preserving redistribution $\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2$ does not change. Also note that our results in the previous section depend on $\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2$. So, the optimal values of the variables mentioned in the previous section do not change at all, which implies that indirect utility of the informal residents does not change at all. In the formal sector, the price of the two goods has not changed but income has changed, so we can say that the rich in the formal sector are better-off and the poor in the formal sector are worse-off because of the widening gap.

Case 2: Now, let us consider the case where either I_1/I_3 increases or I_2/I_3 increases. The analysis would yield similar results for both of them, so let us just consider the case where I_1/I_3 and the change is such that the mean income remains the same. Again, if we look at the optimal values in the previous section, we see that p_f^* increases, q_s^* decreases and N_s^* increases. Therefore, due to rise in the price of housing in the formal sector, the middle class residing in the formal sector are affected negatively although their income remains same. The indirect utility of the informal residents is

$$V_s = [(1-\alpha)I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

Hence the poor in the informal sector are also affected negatively due to lower income and lower consumption of the non-housing good.

The indirect utility of the rich is

$$V_1 = [(1-\alpha)I_1]^{1-\alpha}[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

Note that although the rich have higher income, they also face higher housing prices but despite that their utility will increase (shown in section 4.2).

The above observations can be summarized in the following proposition.

Proposition 2.1 *a) Increase (decrease) in income inequality in the formal sector does not affect the squatter population, however, it negatively (positively) affects the middle income group.*

b) Increase (decrease) in income inequality between the formal sector rich and squatter population negatively (positively) affects the squatter population and the middle income group, positively (negatively) affects the rich and size of squatter population increases (decreases).

2.4 Extention of the Model

In this section, we would like to extend the model by bringing in the government which taxes price of housing in the formal sector and the proceeds from the tax are used to supply public goods to the squatter population. Let t denote the tax of housing price in the formal market. Also, let H denote the level of public facilities available in the squatter land and this enters the utility function of the squatter household in the following way

$$U_s = Hx^{1-\alpha}q^{\alpha}$$

The new budget constraint of the formal residents is

$$(1+t)p_f q_f + x = I_i \text{ for } i = 1, 2.$$

Therefore the demand for housing from the formal sector is

$$d_f(p_f) = \bar{N}_{f_1}\frac{\alpha I_1}{(1+t)p_f} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)p_f}.$$

Tax revenues of the government is

$$T(t) = t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}]$$

Note that the tax revenue does not depend of the price of housing in the formal sector.

Let $H(T(t))$ denote the function which transforms tax revenues to public utilities. The squatter organizer while solving the problem takes H in the utility function to be exogenously given. Therefore his problem becomes

$$\begin{aligned}
\max{}_{\{A,q_s,N_s\}} H(I_3 - A)^{1-\alpha} q_s^{\alpha} \\
\text{subject to } L_s + L_f = \bar{L} \\
N_s q_s = L_s \\
\bar{N}_{f_1} \frac{\alpha I_1}{(1+t)p_f} + \bar{N}_{f_2} \frac{\alpha I_2}{(1+t)p_f} = L_f \\
A N_s k \geq p_f
\end{aligned}$$

Solving the problem similarly we get

$$\begin{aligned}
p_f^* &= \frac{2\alpha(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)}{\bar{L}(1+t)} \\
A^* &= \alpha I_3 \\
q_s^* &= \frac{k I_3 \bar{L}^2 (1+t)}{4(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)} \\
N_s^* &= \frac{2(\bar{N}_{f_1} I_1 + \bar{N}_{f_2} I_2)}{k I_3 \bar{L}(1+t)} \\
L_s^* &= L_f^* = \bar{L}/2
\end{aligned}$$

2.4.1 Effects of Increase in Income Inequality

The analysis in the section will be similar to the analysis in section 3. We want to see how the different groups of people are affected by the increase in income inequality. We will divide our analysis into two cases. Case 1 will discuss the scenario where income inequality rises between the rich and the middle income group and case 2 will discuss the scenario where income inequality increases between the rich and the poor.

Case 1: Suppose I_1 increases and I_2 decreases and also suppose this is a mean preserving spread which again implies as in section 3 that

$$\frac{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2) + 2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)/k\bar{L}}{\bar{N}_{f_1} + \bar{N}_{f_2} + 2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)/kI_3\bar{L}}$$

does not change. Therefore, as before, note that I_1 increases and I_2 decreases but since this is a mean preserving redistribution $\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2$ does not change. Also note that our results in the previous section depend on $\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2$. So, we can conclude that in the presence of a proportional tax t, rise in income inequality affects the rich and the middle income group, but has no affect of the informal settlements, since the total income in the hands of the formal residents has not changed. However, it should be noted that rarely, we have a tax structure which imposes the same tax t for all income levels. If we have a progressive tax structure in place, then in equilibrium, the poor would benefit from an increase in income inequality in the formal sector (if we assume that the function H is increasing).

Case 2: Suppose I_1 increases and I_3 decreases, and again we are considering a mean preserving spread. We can consider the other situation where I_2 increases, I_3 decreases and I_1 does not change but the analysis will be similar in both cases and so we will concentrate on the former case. One can observe from the solution obtained in section 4, that p_f^* goes up, q_s^* goes down and N_s^* goes up. Therefore the indirect utility of the middle income group goes down for sure although their income did not change. The indirect utility of the rich in the presence of tax is

$$V_1 = [(1-\alpha)I_1]^{1-\alpha}[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

and one can show that this function is increasing in I_1.

The indirect utility of the poor is

$$V_s = H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

Now, we would like to compare the indirect utility of the poor before and after increase in inequality. Suppose to start with the income of the groups were $(\bar{I}_1, \bar{I}_3)$ and after the change in income, taking care of the fact that it is a mean preserving spread, let them be $(\beta\bar{I}_1, \delta\bar{I}_3)$, such that $\beta > 1$ and $0 < \delta < 1$. Let $\bar{V}_s$ denote the indirect utility after the change in income inequality. The indirect utility becomes

$$\bar{V}_s = H(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

Therefore, we can say that the squatter population is worse off due to an increase in inequality if and only if

$$H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}]) \geq H(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])\delta(\frac{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)})^{\alpha}$$

Note that the tax revenue has gone up due to the increase in inequality but the poor also has less income compared to before. So it is not clear whether the increase in income inequality between the rich and the poor makes the poor worse off or not. Note that if the population of poor relative to the rich is very high then β would be much larger than δ which would increase the probability that the poor are better off after the increase in income inequality.

Proposition 2.2 *In the presence of government provision of public goods*

a) Increase (decrease) in income inequality in the formal sector does not affect the squatter population, however, it negatively (positively) affects the middle income group.

b) Increase (decrease) in income inequality between the formal sector rich and squatter population negatively (positively) affects the middle income group, positively (negatively) affects the rich and the effect on a squatter household is ambiguous.

This is an important result because it shows that the notion that the poor are always negatively affected due to increase in income inequality might not be true. It depends on the function H and the population size of each group. Note that even if we are in a win-win situation for both rich and the poor, the middle income group suffers unambiguously.

In the following section, we want to see that if we are in a situation where the poor are indeed affected negatively due to a rise in income inequality, then is it possible for the government to increase the tax rate t on formal housing and make the poor better off.

2.4.2 Increase in Tax

In this section, we want to see if rising the tax on formal housing can offset the negative impact of rising income inequality. Let us assume that after the rise in income inequality between the rich and the poor the tax rate t has been increased from t to t'. Therefore as derived in the previous section the squatter population is better off after tax cum inequality change if and only if

$$\begin{aligned}
\bar{V}_s &= H(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}])[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\bar{L}^2(1+t')}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
&\geq H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}
\end{aligned}$$

which simplifying to

$$\bar{V}_s = H(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}])\delta(\frac{1+t'}{1+t})^{\alpha}(\frac{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)})^{\alpha}$$
$$\geq H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])$$

Note that $\delta(\frac{(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}{(\bar{N}_{f_1}\beta I_1+\bar{N}_{f_2}I_2)})^{\alpha} < 1$ so a necessary condition for the poor to be better off is

$$H(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}])(\frac{1+t'}{1+t})^{\alpha} \geq H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])$$

which clearly holds. The objective of doing this is to show that by increasing the tax rate there is a scope for the government to assure that the welfare of a poor household is not negatively affected.

Proposition 2.3 *In the presence of government provision of public goods, higher tax on housing in the formal sector can offset the negative impacts of rising income inequality on the poor households.*

For a linear function of $H(.)$ one can calculate the exact value of t' and that has been done in the appendix. Note that we can further analyze the effect of this on the welfare of the other two income groups. Suppose after the increase in income inequality, the tax rate has been increased from t to t' and also assume, for the sake of simplicity that the function $H(.)$ is linear. We know that at the tax rate t' the poor are just as well off as they were before the rise of income inequality. One can show that the indirect utility of the rich and the middle income group does not change with the introduction of the tax because the price adjusts when the tax rate is increased. The owner of the land who get the rents (p_f) are negatively affected by the increase in tax.

If we divert our attention to the other variables of interest, then there are some interesting things happening as well. Note that when the tax was t, q_s and N_s were

$$
\begin{aligned}
q_s^* &= \frac{k\delta I_3 \bar{L}^2(1+t)}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2} I_2)} \\
N_s^* &= \frac{2(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2} I_2)}{k\delta I_3 \bar{L}(1+t)}
\end{aligned}
$$

After the increase in taxes from t to t' q_s and N_s were

$$
\begin{aligned}
q_s^* &= \frac{k\delta I_3 \bar{L}^2(1+t')}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2} I_2)} \\
N_s^* &= \frac{2(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2} I_2)}{k\delta I_3 \bar{L}(1+t')}
\end{aligned}
$$

So note that although the higher tax rate t' ensured that an individual squatter household is as least as good as before the change in income inequality, we see that the number of squatter households that the city can accommodate decreases, which implies that the cost of bringing this change is quite high. This can be summarized in the following proposition.

Proposition 2.4 *In the presence of government provision of public goods, higher tax lead to higher consumption of housing by the poor but the number of poor people the city can accommodate decreases.*

2.5 Movement of Households Between Sectors

In this section, we want to modify the model by allowing for many income levels in the model. So far, we had three groups of individuals, the rich, the middle-income ones and the poor. The population size of the rich and the poor was exogenously fixed and the population size of the squatter population was endogenously determined in the model. We want to see if there are movements of households from the formal to the informal sector due to rise in income inequality. Therefore, one would think that a model with continuous levels of income of households would be ideal for the analysis and we can get a threshold value of income which determines whether a household chooses to live in the formal sector or the informal sector. If we see that the threshold of income changes with change in income inequality, then we can claim that income inequality not only affects the well-being of each group but also affects the size of the informal sector. Note that we have already shown that the size of the informal sector is indeed affected by the rise of income inequality in the previous section but that was due to the choice the squatter organizer made whether to allow more or less poor in the informal sector. There was no movement between the informal and the formal sector and that is precisely the channel we want to explore in this section.

This can be explored within the frameworks of our existing basic set-up. We will move to the continuous framework later on and before getting to the continuous frameworks let us see what happens if I_2 and I_3 are close to each other. In the previous sections, we did not bother to consider whether the middle income groups would choose to live in the slums or not because the underlying assumption was that incomes of middle income group were sufficiently high to avoid movement of the middle income group to the slums. We derived a condition which ensures that the households in the informal sector would not choose to move to the formal sector. The proposition in the previous section showed that when income inequality increased between the poor and the rich the middle income group was negatively affected. In what follows we will see under what conditions would the middle income group be willing to move to the informal sector.

The analysis can be divided into two parts. First, we will look at the effects of rising income inequality between the rich and the middle income group and then we will move on to the effects of rising income inequality between rich and poor.

Case 1: Suppose income inequality increases between the rich and the middle income group. At income level I_2 the indirect utility of the middle income group was

$$V_m = [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}$$

Suppose the income of the rich goes up from I_1 to θI_1 and that of the middle income group goes down from I_2 to μI_3 such that $\theta > 1$ and $\mu < 1$. Note that for this to be a mean preserving spread the following condition should hold

$$\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2 = \bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}\mu I_2$$

The indirect utility of the middle income group now becomes

$$\begin{aligned} V'_m &= [(1-\alpha)\mu I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}\mu I_2)}]^{\alpha} \\ &= [(1-\alpha)\mu I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \end{aligned}$$

Clearly utility of the middle income group goes down. If, for the sake of simplicity, we assume that all slum dwellers have to pay the same for defensive activities, and gets the same amount of housing, then middle income group would want to reside in the informal area if and only if

$$\begin{aligned} H(t[\bar{N}_{f_1}\frac{\alpha\theta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha\mu I_2}{(1+t)}])[\mu I_2 - \alpha I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\ \geq [(1-\alpha)\mu I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}, \end{aligned}$$

which implies

$$\begin{aligned} H(t[\bar{N}_{f_1}\frac{\alpha\theta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha\mu I_2}{(1+t)}])[\mu I_2 - \alpha I_3]^{1-\alpha}[kI_3\bar{L}(1+t)/2]^{\alpha} \\ \geq [(1-\alpha)I_2]^{1-\alpha}\mu I_2^{\alpha} \end{aligned}$$

Therefore we see that rise in income inequality can potentially cause movements of households from the formal to the informal sector. Note that till now we have not solved for the equilibrium of any sort in this case and we will do that later but even without solving one can see that there are incentives for the middle income group to move to the slums.

Case 2: Suppose the income inequality rises between the rich and the poor and the middle income group's income remains the same. As we did previously, let θ denote the rise in income of the rich and β denote the fall in income of the poor and θ and β are such that the redistribution is mean preserving. The indirect utility of the middle income become rise in income inequality was

$$V_m = [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}.$$

After the rise in income inequality the indirect utility of the middle income household is

$$V'_m = [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}.$$

Therefore note that again there is a drop in the utility of the middle income group due to rise in income inequality. The middle income group would like to move to the informal sector if and only if

$$\begin{aligned} H(t[\bar{N}_{f_1}\frac{\alpha\theta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[I_2 - \alpha I_3]^{1-\alpha}[\frac{k\beta I_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\ \geq [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}, \end{aligned}$$

As we had seen in the previous case, the rise in income inequality might induce middle income group to move into the informal sector. Also, note that we can compare the two cases to see which situation makes the middle income group relatively worse. The middle income group becomes worse in the situation of case 1 than case 2 if and only if

$$\begin{aligned} [(1-\alpha)\mu I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} &\leq [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\ \mu &\leq \frac{\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2}{\bar{N}_{f_1}\theta I_1 + \bar{N}_{f_2}I_2}. \end{aligned}$$

So note that it is not clear whether the middle income households are worse off when their income decreases versus when the poor become poorer. These observations are summarized in the following proposition.

Proposition 2.5 *a) Rise in income inequality between rich and middle income and between rich and poor can result in movement of middle income households from formal sector to the informal sector.*
b) Also under certain conditions rise in income inequality between rich and poor can make middle income households in the formal sector worse than if there own income had gone down.

Note that in the previous analysis we have considered cases where income inequality only rises but one can also so a similar analysis for analyzing situations when income inequality falls. Suppose income inequality decreases between the rich and the middle income groups. It is easy to see that if it is a mean-preserving spread, then that would not affect the poor. But instead, suppose we are in a situation where rich and middle income live in the formal sector and middle income and poor live in the informal sector. Then fall in income inequality can potentially induce the middle income households in the informal sector to move back to the formal sector. The motive of this discussion is to point out that changes in income inequality between different groups of households change the equilibrium of the city significantly. The problem of growing slum population can be significantly intensified by rising income inequality. Therefore it is necessary for the policy makers to consider the sources of growing slum population and also address them.

Till now we have tried to lay down conditions under which people would want to move across sectors and our focus has been the middle income group. We have shown that movement of middle income groups from formal to informal sector is possible in both situations- when income inequality rises between rich and poor or when income inequality rises between rich and middle income groups. Now we will try to formulate the problem differently and find the equilibrium under the assumption that there is an incentive for the middle income group to move. The difficult part of this problem is to model the squatter organizer's problem and one can take different directions in terms of modeling. One way of doing it is suppose the squatter organizer maximizes the sum of utilities of the poor and the middle income group who wants to stay in the informal sector. Then the squatter organizer would maximize

$$H[N'_{f_2}(\mu I_2 - A_2)^{1-\alpha} q_{sm}^{\alpha} + N_s(I_3 - A_3)^{1-\alpha} q_{sp}^{\alpha}]$$

where N'_{f_2} denotes the number of middle income households in the informal sector, q_{sm} and q_{sp} denote the housing consumption by the middle income household and the poor household in the informal sector and A_2 aand A_3 denote the contribution for prevention of eviction in the squatter land. The organiser would choose A_2, A_3, N_s, q_{sm} and q_{sp}. The constraints faced by the organiser is as follows

$$\begin{aligned}
L_s + L_f &= \bar{L} \\
N_s q_{sp} + N'_{f_2} q_{sm} &= L_s \\
\bar{N}_{f_1} \frac{\alpha \theta I_1}{(1+t)p_f} + (\bar{N}_{f_2} - N'_{f_2}) \frac{\alpha \mu I_2}{(1+t)p_f} &= L_f \\
k[A_3 N_s + A_2 N'_{f_2}] &\geq p_f \\
(\mu I_2 - A_2)^{1-\alpha} q_{sm}^{\alpha} &= [(1-\alpha)I_2]^{1-\alpha} [\frac{\bar{L} \mu I_2}{2(\bar{N}_{f_1} \theta I_1 + (\bar{N}_{f_2} - N'_{f_2}) \mu I_2)}]^{\alpha}.
\end{aligned}$$

Note that we are considering the situation under which income inequality has gone up between the rich and the middle income housholds and the last constraint ensures that at the equilibrium no more middle income group houshold finds it beneficial to move from the formal to the informal sector. If instead we were in the situation where income inequality had increased between the rich and the poor then the squatter organiser's problem would have been as follows.

$$
\begin{aligned}
\max H[N'_{f_2}(I_2 - A_2)^{1-\alpha} q_{sm}^{\alpha} \quad &+ \quad N_s(\beta I_3 - A_3)^{1-\alpha} q_{sp}^{\alpha}] \\
\text{subject to} & \\
L_s + L_f \quad &= \quad \bar{L} \\
N_s q_{sp} + N'_{f_2} q_{sm} \quad &= \quad L_s \\
\bar{N}_{f_1} \frac{\alpha \theta I_1}{(1+t)p_f} + (\bar{N}_{f_2} - N'_{f_2}) \frac{\alpha I_2}{(1+t)p_f} \quad &= \quad L_f \\
k[A_3 N_s + A_2 N'_{f_2}] \quad &\geq \quad p_f \\
(I_2 - A_2)^{1-\alpha} q_{sm}^{\alpha} \quad &= \quad [(1-\alpha)I_2]^{1-\alpha} [\frac{\bar{L} I_2}{2(\bar{N}_{f_1} \theta I_1 + (\bar{N}_{f_2} - N'_{f_2}) I_2)}]^{\alpha}.
\end{aligned}
$$

Alternatively, one can choose to model in the following way where the squatter organizer cares only about the poor population living in the informal sector and everybody living in the informal sector gets the same amount of housing and pays the same A irrespective of their income. Therefore the problem faced by the squatter organizer is as follows

$$\max(I_3 - A)^{1-\alpha} q_s^{\alpha}]$$

subject to

$$\begin{aligned}
L_s + L_f &= \bar{L} \\
(N_s + N'_{f_2})q_s &= L_s \\
\bar{N}_{f_1}\frac{\alpha\theta I_1}{(1+t)p_f} + (\bar{N}_{f_2} - N'_{f_2})\frac{\alpha\mu I_2}{(1+t)p_f} &= L_f \\
k[A(N_s + N'_{f_2})] &\geq p_f \\
(\mu I_2 - A)^{1-\alpha} q_s^{\alpha} &= [(1-\alpha)\mu I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}\theta I_1 + (\bar{N}_{f_2} - N'_{f_2})\mu I_2)}]^{\alpha}.
\end{aligned}$$

This is the case where income inequality increased between the rich and the middle income hosuholds. The organiser chooses N_s, A and q_s. In a similar manner one can formulate the problem of the organiser when income inequality rises between rich and poor.

$$\max(\beta I_3 - A)^{1-\alpha} q_s^{\alpha}]$$

subject to

$$\begin{aligned}
L_s + L_f &= \bar{L} \\
(N_s + N'_{f_2})q_s &= L_s \\
\bar{N}_{f_1}\frac{\alpha\theta I_1}{(1+t)p_f} + (\bar{N}_{f_2} - N'_{f_2})\frac{\alpha I_2}{(1+t)p_f} &= L_f \\
k[A(N_s + N'_{f_2})] &\geq p_f \\
(I_2 - A)^{1-\alpha} q_s^{\alpha} &= [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}\mu I_2}{2(\bar{N}_{f_1}\theta I_1 + (\bar{N}_{f_2} - N'_{f_2})I_2)}]^{\alpha}.
\end{aligned}$$

2.5.1 Movement of Poor From Informal to Formal Land

It is often discussed in the policy circle that despite inequality rising in the economy, poverty has decreased a lot. The following figure shows the decrease in poverty levels in both rural and urban India.

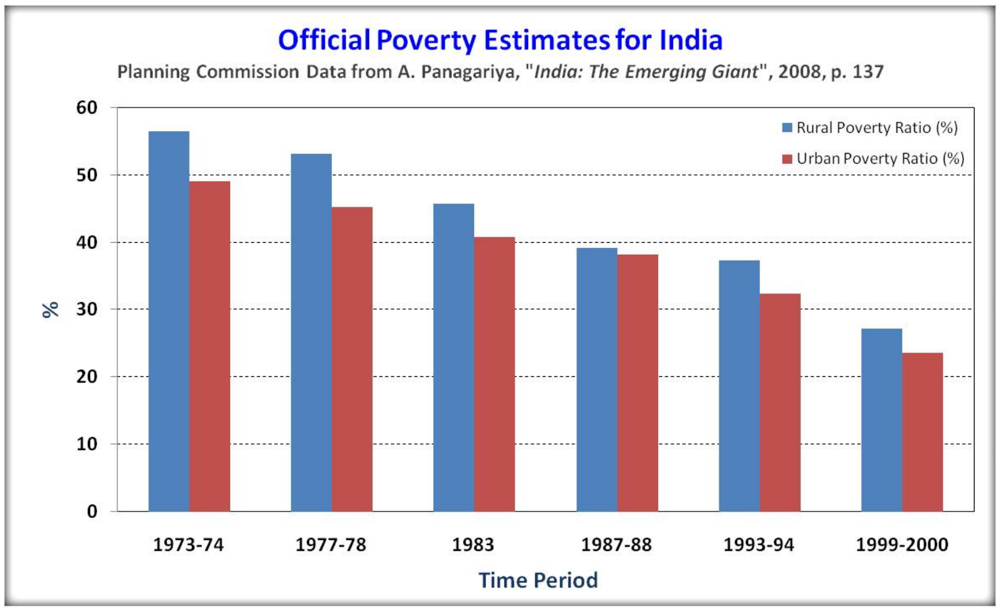


Figure 2.3:

Therefore if it is the case that poverty is going down and income inequality is going up, then one can analyse the situation where everyone has more income but the rate as which the rich are becoming righer is greater than the rate as which the income of the poor are increasing. For the sake of simplicity, we will assume that a $\bar{N}_s$ number of people who had a income of I_3 now have an income I_2. Note that it is not necessary to assume that they move to the middle income group, we can choose some other income below I_2 as well but that unnecessarily complicates the problem with adding much. Also we will asume that I_2 is sufficiently high so that the middle income group does not want to stay in the informal sector. Let us define the rate of increase of income of the $\bar{N}_s$ poor as $I_2 - I_3 = k$. Therefore going by our story I_1 has to increase to something more than $I_1 + k$, and let it be denoted by I_1'. Therefore we can rewrite the planner's problem as follows

$$\begin{aligned}
\max{}_{\{A,q_s,N_s\}} H(I_3 - A)^{1-\alpha} q_s^{\alpha} \\
\text{subject to } L_s + L_f = \bar{L} \\
N_s q_s = L_s \\
\bar{N}_{f_1} \frac{\alpha(I_1 + k)}{(1+t)p_f} + (\bar{N}_{f_2} + \bar{N}_s)\frac{\alpha I_2}{(1+t)p_f} = L_f \\
A N_s k \geq p_f
\end{aligned}$$

Solving the problem similarly we get

$$\begin{aligned}
p_f^* &= \frac{2\alpha(\bar{N}_{f_1}(I_1 + k) + (\bar{N}_{f_2} + \bar{N}_s)I_2)}{\bar{L}(1+t)} \\
A^* &= \alpha I_3 \\
q_s^* &= \frac{k I_3 \bar{L}^2 (1+t)}{4(\bar{N}_{f_1}(I_1 + k) + (\bar{N}_{f_2} + \bar{N}_s)I_2)} \\
N_s^* &= \frac{2(\bar{N}_{f_1}(I_1 + k) + (\bar{N}_{f_2} + \bar{N}_s)I_2)}{k I_3 \bar{L}(1+t)} \\
L_s^* &= L_f^* = \bar{L}/2
\end{aligned}$$

The results show that the price of housing in the formal sector has gone up due to higher demand in the formal sector, the number of people the staying in the informal sector has gone up but their consumption of housing has gone down. So we see that as a result of some poor population becoming richer and increasing the size of the middle income group, there is a rise in the number of slum dwellers. This result is interesting because one would think that higher income of the poor would help them to have better housing and hence would stop the growth in the informal settlement but the constant migration of poor from rural to urban sector does not allow stop the growth of informal settlements, in fact, the opposite happens. Also, a result of the growth in the size of the middle income group, the utility of each household in the middle income group also decreases, due to rise in the price of land in the formal sector. It can be seen easily from the indirect utility of the middle income group.

$$V_2 = [(1-\alpha)I_2]^{1-\alpha}[\frac{\bar{L}I_2}{2(\bar{N}_{f_1}(I_1+k)+(\bar{N}_{f_2}+\bar{N}_s)I_2)}]^{\alpha}.$$

Tax collected after the change in incomes is as follows

$$T(t) = t[\bar{N}_{f_1}\frac{\alpha(I_1+k)}{(1+t)} + (\bar{N}_{f_2}+\bar{N}_s)\frac{\alpha I_2}{(1+t)}],$$

which shows that tax collected has increased because of higher income of rich and more number of middle income households. This implies that it is ambiguous whether poor living in the informal land are better off or worse off after the change in inequality. All the above observations is summarized in the following proposition.

Proposition 2.6 *If income increases for both rich and poor and income inequality rise, then the rich are better off, the original middle income group are worse off, the effect on informal settlers is ambiguous but the number of informal settlers increases.*

2.6 Effects of Expansion of City Limits

By expansion of city limits, we mean an increase in $\bar{L}$. It is commonly observed that big cities in developing countries are always expanding. One potential reason might be that the government is building roads, or expanding transport facilities which allow individuals to reside in those places thus increasing the size of the city. Simple partial derivatives of the optimal values obtained in section 5 show that expansion of city limit will result in a lower price of housing in the formal sector, non-linear increase in housing consumption of the squatter population and decrease in size of the squatter population. Both the rich and the middle income group are unambiguously better off and the utility of the poor after the increase in land size from *barL* to $\tilde{L}$ is

$$V_s = H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)I_3]^{\alpha}[\frac{kI_3\tilde{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}.$$

Note that the utility of the poor also goes up unambiguously. This also suggests that if a city is expanding then it is possible for the city counter the negative impacts of rising income inequality through the expansion of land size. Note that if there is rising income inequality and rise in land size simultaneously then the utility of the poor will be

$$\bar{V}_s = H(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\tilde{L}^2(1+t)}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}.$$

Therefore the condition which has to hold for the poor to be not worse off after the rise in inequality and land size is

$$\begin{aligned} &H(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\tilde{L}^2(1+t)}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\ &\geq H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)I_3]^{\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}. \end{aligned}$$

2.6.1 A Linear H Function

Suppose we have a linear H function which takes the following form

$$H = a + bk;\ a, b > 0.$$

Then the condition for the poor not to be worse off is

$$\begin{aligned}
& [a + b(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])][(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\tilde{L}^2(1+t)}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
\geq\ & [a + b(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])][(1-\alpha)I_3]^{\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
\Longrightarrow\ & [a + b(t[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])]\delta[\frac{\tilde{L}^2}{\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2}]^{\alpha} \\
\geq\ & [a + b(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])][\frac{\bar{L}^2}{\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2}]^{\alpha}
\end{aligned}$$

Proposition 2.7 *Under certain conditions expansion of city limit can undo the negative impact of rising income inequality between rich and poor.*

These results are driven by the fact that the population in the formal sector is fixed and since more land is available now the price of land goes down. The consumption of housing of each squatter households grows non-linearly which prevents other poor people entry into the squatter land.

To emphasize the fact that the assumption of fixed formal sector residents is crucial, let us do a small exercise. Suppose land increases from $\bar{L}$ to $r\bar{L}$ such that $r > 1$ and both groups of formal residents increase λ times such that $\lambda > 1$. Solving the planner's problem again one would get

$$
\begin{aligned}
p_f^* &= (\frac{\lambda}{r})\frac{2\alpha(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{\bar{L}(1+t)} \\
A^* &= \alpha I_3 \\
q_s^* &= (\frac{r^2}{\lambda})\frac{(1+t)kI_3\bar{L}^2}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)} \\
N_s^* &= (\frac{\lambda}{r})\frac{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{kI_3\bar{L}(1+t)} \\
L_s^* &= L_f^* = \bar{L}/2
\end{aligned}
$$

Therefore, one can see that if the expansion of city limits is accompanied by the influx of formal residents into the city, then the above mentioned observations might not hold. The results will depend on the relative increment size of land and of the size of formal residents.

2.7 Policy Discussions

Several governments is many countries had to face or are still facing this problem of rising informal settlements and various government policies have been tried in different countries to tackle this problem. In India, the projections of future informal settlements are quite alarming and hence the Indian goverment has set a new target called ”Housing for All by 2022”. Most policies, till date, aimed at reducing informal population, can be classified into two broad categories:

i) Uprooting informal settlements completely by brute force since informal settlers do not have any legal right on the land. Policy makers quickly realized that this has a high social cost and hence was not used too much.

ii) Slum redevelopment and better supply of basic ammenities. These policies aimed at helping informal households to improve their existing houses and also provide basic ammenities like clean water and electricity. "Housing for All" in India is a different version of slum redevelopment process where they also aim to have a public-private partnership in building new houses for the informal settlers and has a target of 300 Indian cities by 2019.

Usually all these policies are long term and take many years to acheive the goal. Every government aim to to bring the whole society into the formal housing section or in other words formalize the informal land in some way. Brueckner and Selod (2009) discuss the consequences of complete formalization of the city space and show that the rich are better off and the poor worse off. In this section, we would discuss the consequences of partial formalization of land. Note that any policy cannot formalize a city overnight, so formalization happens in phases. If that is the case then it would be interesting to do a small exercise on partial formalization and its consequences.

To begin with, let the take the simplest model we cosidered in the beginning of the paper with government and no provision of public goods. The way we want to think about partial formalization is that the goverment chooses a plot of land which was previously occupied by informal settlers and builds housing and accomodates as many informal settlers as in can. Let us assume ϕ fraction of available land in the city is where the formalization has taken place. So, the city is left with $(1-\phi)\bar{L}$ of land area. Setting up the problem in a similar manner to before, the squatter organizer faces the following problem:

$$\begin{aligned}
\max{}_{\{A,q_s,N_s\}}(I_3 - A)^{1-\alpha} q_s^{\alpha} \\
\text{subject to } L_s + L_f = \bar{L}(1-\phi) \\
N_s q_s = L_s \\
\bar{N}_{f_1}[\frac{\alpha I_1}{p_f}] + \bar{N}_{f_2}[\frac{\alpha I_2}{p_f}] = L_f \\
A N_s k \geq p_f
\end{aligned}$$

Rearranging equations in the problem as before, the problem can be rewritten as (planner chooses N_s and A)

$$\max\ (I_3 - A)^{1-\alpha}(\frac{\bar{L}(1-\phi) - \beta/AN_s}{N_s})^{\alpha} \tag{2.2}$$

where $\beta = \frac{(\bar{N}_{f_1}\alpha I_1 + \bar{N}_{f_2}\alpha I_2)}{k}$.

Solving the problem would yield the following solution

$$\begin{aligned}
p_f^* &= \frac{2\alpha(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{\bar{L}(1-\phi)} \\
A^* &= \alpha I_3 \\
q_s^* &= \frac{kI_3[\bar{L}(1-\phi)]^2}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)} \\
N_s^* &= \frac{2(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}{kI_3\bar{L}}(1-\phi) \\
L_s^* &= L_f^* = \bar{L}(1-\phi)/2
\end{aligned}$$

Proposition 2.8 *Under partial formalization, all sections of the society are negatively affected, the formal sector faces higher rent and the informal sector gets less units of housing per household and the number of informal settlers increases.*

Note that because land becomes scarce, the price of land in the formal sector goes up which negatively affects both the rich and the middle income groups. Since $\hat{p}_f$ is higher now, the squatter organizer has to accomodate more people in the informal sector to satisfy the no eviction constraint. This leads to a lower q_s and hence affects the informal settlers negatively.

We can move further and conduct an analysis to check how increasing income inequality affects different sections of the society when the society in in a state 1of partial formalization. We believe that it would be interesting and we might see changes in magnitudes of the effect but the qualitative results would not be affected.

This analysis on partial formalization also suggests that it is also important to expand city limits if possible, as this would help in reducing the negative impacts poited out in the previous proposition. We understand that expanding cities is not so easy in a country like India where the city expansion rate is very low. Also, since the migration rates to cities are quite high, this suggests that government policies to aim towards vertical growth of cities. This also suggests that redevelopment of existing informal housing might not be a good policy instrument in the long run.

2.8 Conclusion

The rising slum population in developing countries is a potentially big problem for most of the developing countries. The objective of this paper is to see if there are other forces with distort the equilibrium of a city and participates in the increase or decrease in slum population. Many countries across the globe are experiencing a rise in income inequality over the last few decades. So the key question was whether rising income inequality plays a role in increasing or decreasing slum population and how that changes the equilibrium of the city and the welfare of nonpoor sections of the society.

In order to do that we relied on a simple theoretical set-up inspired from Brueckner and Selod's work. We modify the model to a large extent to allow for the governmental provision of public goods in the informal sector and also extend the model to allow for two groups of households residing in the formal sector- the rich and the middle income group. We assume that government taxes the housing sector in the formal sector for funding the provision of public good and it balances the budget. The results shows some interesting features. We see that the results crucially depend on the whether income inequality rises between rich and middle income groups or between rich/middle income and poor households. Our preliminary results suggest that rise in income inequality between the rich and middle income gruops does not have any effect on the slum dwellers or the informal settlers. However, rise in income inequality between the poor and the rich affects the poor negatively and interestingly it also affects the middle income group. Then we moved on to introduce government provision of public good in the informal sector. Results show that as before, rise in income inequality between rich and middle income groups does not affect the poor as long as the average income remains the same. But when income inequality rises between the rich and the poor it is not clear whether that hurts the poor or not. It crucially depends on the rise in government provision of public goods and the extent to which the poor households income goes down. This shows that firstly if income inequality as a whole goes up or down, it is important to find out which sections of the society suffered and gained and secondly to note that it might not be always bad for the poor if income inequality goes up. We went on to show that conditions under which the poor are indeed negatively affected due to a rise in income inequality between the rich and the poor, a higher tax on housing can reduce the negative impact of rising inequality and there exists a tax rate t which makes the poor just as well off as it was before. Finally, we considered the issue of movement of households across sectors. We show that if the income of the middle income group is sufficiently low, then under both cases (rise in inequality between rich and poor and rise in inequality between rich and poor) there is a possibility that middle income group households would want to reside in the informal sector. We address a common concern related to decrease in poverty and rise in income inequality. Many countries have witnessed large decrease in poverty levels but huge rise in income inequality. This is possible if both poor and rich become richer, but at differential rates. We find that in this situation the number of poor households in the city increases; the effect on their welfare is ambiguous but the middle income group is worse off.

We also look at the issue of partial formalization of slums and find interesting results and challenges for the policy makers. We saw that partial formalization will negatively affect all sections of the society. Our analysis suggests that with lack of city expansion, vertical expansion of cities is the only way to accomodate huge population in cities and therefore redevelopment of existing informal housing might not be a good instrument in the long run.

Lack of data on these issues makes us rely on theoretical models to discuss the issue. In a simple set-up which models informal settlements explicitly, we see that rise or fall in income inequality amplifies or reduces the rising growth of slum population in developing countries. Income inequality has the potential to add fuel to the fire and therefore policy makers need to consider these issues while discussing slum redevelopment policies.

2.9 References

1. Alonso, W. (1964). "Location and land use, *Cambridge: Harvard University Press*.

2. Brueckner, J. K. and Selod, H. (2009). "A Theory of Urban Squatting and Land-Tenure Formalization in Developing Countries", *American Economic Journal: Economic Policy*, Volume 1, pp. 28-51.

3. Glazer, A., Kanniainen, V., and Poutvaara, P. (2008). "Income taxes, property values, and migration", *Journal of Public Economics*, Volume 92, pp. 915-923.

4. Gyourko, J., Mayer, C., and Sinai, T. (2013). "Superstar Cities", *American Economic Review: Economic Policy*, Volume 5, pp. 167-99.

5. Landvoigt, T., Piazzesi, M., and Schneider, M. (2015). "The Housing Market(s) of San Diego", *American Economic Review*, Volume 105, pp. 1371-1407

6. Maattanen, N., and Tervi, M. (2014). "Income distribution and housing prices: An assignment model approach", *Journal of Economic Theory*, Volume 151, pp. 381-410.

7. Matlack, J. L., and Vigdor, J. L. (2008) "Do rising tides lift all prices? Income inequality and housing affordability", *Journal of Housing Economics*, Volume 17, pp. 212-224.

8. Mansur, E. T., Quigley, J., Raphael, S., and Smolensky, E. (2002). "Examining policies to reduce homelessness using a general equilibrium model of the housing market", *Journal of Urban Economics*, Volume 52, pp. 316-340.

9. Mills, E.S. (1967), “An aggregative model of resource allocation in a metropolitan area, *American Economic Review*, Volume 57, pp. 197-210.

10. Muth, R.F. (1969), “Cities and housing, *Chicago: University of Chicago Press*.

11. Nieuwerburgh, S. V. and Weill, P. (2010). “Why Has House Price Dispersion Gone Up?”, *The Review of Economic Studies*, 77 (4): 1567-1606.

12. Quigley, J., Raphael, S., and Smolensky, E. (2001). “Homeless In America, Homeless In California”, *The Review of Economics and Statistics*, Volume 83, issue 1, 37-51.

2.10 Appendix

2.10.1 Derivation of Solution of the Squatter Planner's Problem

Differentiating equation (1) with respect to A and N_s yields the following respectively

$$-(1-\alpha)(\bar{L}N_sA^2 - \beta A) + \alpha(I_3 - A)\beta = 0 \tag{2.3}$$

$$\frac{\beta}{AN_s^2} - \frac{(\bar{L} - \beta/AN_s)}{N_s} = 0 \tag{2.4}$$

Rearranging equation (3) yields

$$AN_s = \frac{2\beta}{\bar{L}} \tag{2.5}$$

Using equation (4) in equation (2) yields $A = \alpha Y_3$. One can also check the second order conditions for a maximization problem and those conditions are satisfied in this case.

2.10.2 Derivation of Solution of the Squatter Planner's Problem with tax and H

The problem is

$$
\begin{aligned}
\max_{\{A,q_s,N_s\}} H(I_3 - A)^{1-\alpha} q_s^{\alpha} \\
\text{subject to } L_s + L_f &= \bar{L} \\
N_s q_s &= L_s \\
\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)p_f} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)p_f} &= L_f \\
A N_s k &\geq p_f
\end{aligned}
$$

which can be rewritten as

$$
\begin{aligned}
\max_{\{A,q_s,N_s\}} H(I_3 - A)^{1-\alpha} q_s^{\alpha} \\
\text{subject to } L_s + L_f &= \bar{L}
\end{aligned}
$$

$$
\begin{aligned}
\text{subject to } L_s + L_f &= \bar{L} \\
N_s q_s &= L_s \\
\tilde{N}_{f_1}\frac{\alpha I_1}{p_f} + \tilde{N}_{f_2}\frac{\alpha I_2}{p_f} &= L_f \\
A N_s k &\geq p_f
\end{aligned}
$$

where $\tilde{N}_{f_1} = \frac{\bar{N}_{f_1}}{1+t}$ and $\tilde{N}_{f_2} = \frac{\bar{N}_{f_2}}{1+t}$. Also notice that the planner takes $H(.)$ to be exogenous and therefore that term vanishes when we take the first order conditions. So the problem reduces to the same problem solved in section 8.1 and for getting the optimal values we can replace $\bar{N}_{f_1}$ and $\bar{N}_{f_2}$ by $\tilde{N}_{f_1}$ and $\tilde{N}_{f_2}$ respectively.

2.10.3 Indirect Utility of Rich When (I_1/I_3) Increases

We can take the pratial derivation of V_1 with respect to I_1 and check if it is positive or not.

$$
\begin{aligned}
\frac{\partial V_1}{\partial I_1} &= (1-\alpha)[(1-\alpha)I_1]^{-\alpha}[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}]^{\alpha} \\
&+ [(1-\alpha)I_1]^{1-\alpha}\alpha[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}]^{\alpha-1} \\
&. [\frac{\bar{L}}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)} - \frac{\bar{L}I_1\bar{N}_{f_1}}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)^2}]
\end{aligned}
$$

$$
\begin{aligned}
&= (1-\alpha)[(1-\alpha)I_1]^{-\alpha}[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}]^{\alpha} \\
&+ [(1-\alpha)I_1]^{1-\alpha}\alpha[\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}]^{\alpha-1} \\
&. [\frac{\bar{L}I_1}{2(\bar{N}_{f_1}I_1+\bar{N}_{f_2}I_2)}] > 0
\end{aligned}
$$

2.10.4 Tax rate with a linear $H(.)$

Let us assume that the function $H(.)$ is as follows

$$H(k) = a + bk;\ a, b > 0,\ b > 1$$

Then the poor will benefit from an increase in tax after an increase in income inequality if and only if the following condition holds

$$
\begin{aligned}
& H(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}])[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\bar{L}^2(1+t')}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
\geq \quad & H(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}])[(1-\alpha)I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha}
\end{aligned}
$$

Using the function H mentioned above, this condition boils down to

$$
\begin{aligned}
& (a + b(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}]))[(1-\alpha)\delta I_3]^{1-\alpha}[\frac{k\delta I_3\bar{L}^2(1+t')}{4(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
\geq \quad & (a + b(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}]))[(1-\alpha)I_3]^{1-\alpha}[\frac{kI_3\bar{L}^2(1+t)}{4(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)}]^{\alpha} \\
\Longrightarrow \quad & (a + b(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}]))\delta(\frac{1+t'}{(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)})^{\alpha} \\
\geq \quad & (a + b(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}]))(\frac{1+t}{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)})^{\alpha} \\
\Longrightarrow \quad & (\frac{1+t'}{1+t})^{\alpha} \geq \frac{((a + b(t[\bar{N}_{f_1}\frac{\alpha I_1}{(1+t)} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t)}]))}{((a + b(t'[\bar{N}_{f_1}\frac{\alpha\beta I_1}{(1+t')} + \bar{N}_{f_2}\frac{\alpha I_2}{(1+t')}]))}(\frac{(\bar{N}_{f_1}\beta I_1 + \bar{N}_{f_2}I_2)}{(\bar{N}_{f_1}I_1 + \bar{N}_{f_2}I_2)})^{\alpha}\frac{1}{\delta}
\end{aligned}
$$

CHAPTER 3

The Effects of Construction Cost Differentials on City Equilibrium in a Monocentric City Set Up

3.1 Introduction

In the world of rapid urbanization, the number of cities around the globe is rapidly increasing. But with rising number of cities and the constant influx of population into the cities, the governments of the respective countries face problems in several dimensions. In most cases the cities are not well equipped in terms of infrastructure facilities like roads, housing, access to public facilities etc, which leave different income groups in the population with disproportionate levels of standard of living. Even if one accounts for the income differences among individuals, a new city has differential costs of construction at different places in the city which also creates additional hurdles in the development of the city. If one looks at two big cities like Rio De Jenerio (Brazil) and Mumbai (India), historical evidence suggests that slums developed in areas where construction costs were high (for example the hills in Rio or the marsh lands in Mumbai) and therefore the poor people occupied those lands and lived under miserable conditions. In Rio De Jenerio, the favela developed on the hilly part of Rio in the late 19th century when homeless poor individuals started living in those places. As pointed out by the Brazilian Institute of Geography and Statistics (IBGE) in 2010, that about 6 percent of the Brazilian population lived in slums, which implies that 11.4 million of the 190 million people that lived in the country resided in areas lacking the basic amenities and were disproportionately exposed to environmental disaster. Similar situations gave rise to slums in Mumbai as well during the early 19th century under the colonial rule. Slums like Dharavi (one of the largest in the world) grew up in the marsh lands of Mumbai as the colonial rulers did not deem it fit for their stay and left it for the poor immigrants and locals.

According to UN-HABITAT, around 33% of the urban population in the developing world in 2012, or about 863 million people, lived in slums. Census data projections show that India's slum population will surge to 104 million by 2017 - or around 9% of the total projected national population of 1.28 billion that year. Census data also shows how a large section of the urban population lives in slums in the four metropolitan cities – as high as 41.3% in Greater Mumbai, 29.6% in Kolkata, 28% in Chennai and about 15% in Delhi. Therefore, coupled with the fact that economic development of developing countries leads to higher urban population and to the emergence of new cities, it also leads to higher slum population in the respective countries.

There is an extensive literature which tries to point out the various reasons and ways in which the population is segregated in the city. Segregation might be caused by differences in income, changes in the transportation technology, rise in income inequality and many others. There are multiple papers which use the monocentric city model and explore how changes in income, transportation technology change the equilibrium of the city. In this paper, we explicitly model housing suppliers and show that it might not be the case that households with the highest bid-rent at any distance from the Central Business District get to live on that place. This is where we deviate from the literature and show that differences in quality of housing and differences in cost of production of those qualities yield a different equilibrium to what we would have got in the absence of those features. This might help us in explaining why certain sections of the population live in certain areas of the city.

3.2 Literature Review

There is an extensive literature which attempts to study the formation of segregated neighborhoods and related policies. Hoff and Sen (2005) shows how in the presence of the private provision of public good, individuals sort themselves into different communities according to income. On the other hand, Brueckner and Selod (2009) considers informal settlements as those which are not paying shows and shows that policies aimed at complete formalization of the city might affect the rich or the previously formal settlers negatively. Hoy and Jimenez (1991) proposes a theoretical model which predicts commonly observed squatter characteristics like landowners not collecting rents from squatters etc. Turnbull (2004) shows that the occurrence of squatter population can be explained by landowner's decision to not exercise property rights and not so much by incomplete land markets. Brueckner (2013) and Shah (2013) extends this literature by considering rent-seeking squatter organizers and squatting in free government land.

We also have the hedonic style models initiated by Rosen (1974) which has been widely used in the literature. The basic structure of these class of model emphasizes that things like price of housing is not just the price of the house, but any price of house observed in reality is a composite index of several prices and the function which links several prices to the composite price may be non-linear. Moreover there is empirical evidence supporting this view. The other class of models is the monocentric city models which was developed in the 1960's and 1970's (see Alonso (1964), Mills (1972) and Muth (1969)). The major criticisms to the monocentric city models is the argument that cities are polycentric. Well, this cannot be denied, but one can also not deny that monocentric city is a fair assumption for most cities around the globe. LeRoy and Sonstelie (1983) shows that the transportation technology can change a city's dynamics completely where by the rich who can afford cars can choose to stay in the outskirts of the city. Bertaud and Brueckner (2003) analysis building height restriction in a closed monocentric city set-up.

Coming to the policy issues of slum redevelopment Bento etal (2008) tries to analyze a policy of improving housing in situ versus relocation. Recently Barnhardt etal (2015) shows that if even slum dwellers are given the opportunity to move to better locations or better houses, they might not choose to because of neighborhood considerations and even those who choose to move might be worse off due to destruction of social capital and this points out that simple policies of relocation might not work.

3.3 Model

In this section, we will discuss the basic assumptions in the model. We have primarily two sides in the model, the demand side and the supply side. The demand side is comprised of housing consumers of two types, households who are rich and households who are poor and the supply side is comprised of two types of suppliers, one which specializes in production of high quality housing and one which specializes in production of low quality housing. The following two sections discusses the assumptions and implications of the demand side and supply side of our model.

3.3.1 Demand Side

We consider a simple model with two groups of households in an open monocentric city model. We have rich households with income w_H and poor households with income w_L. Since this is an open city model, how many households of each type, the city will accommodate will be determined when we solve the model. We also assume that the size or quantity of a house is fixed. Housing differs only in one aspect and that is housing quality. Housing quality can be of two types: high quality housing denoted by (q_H) and low quality housing denoted by (q_L). To simplify our model, we further assume that rich households consume high quality housing and poor households consume low quality housing. It is important to note that the model would not change a lot if we do not assume this and allow both types of households to consume both types of housing.

All city residents travel to the Central Business District (CBD) for work. There are two modes of transportation available in the city: the bus denoted by (b) and the car denoted by (a). Households can choose to travel by public transportation which is the bus or can choose to travel by private transportation which is the car.

Travelling cost: -

- Car: $f^a + v^a x + w_i t^a x$
- Bus: $f^b + v^b x + w_i t^b x$ for $= H, L$.

x denote the distance a household lives away from the CBD and hence that is the distance it must travel to work. There are three types of costs involved in travelling. We have the fixed cost of choosing any mode of transport (denoted by f), the variable cost of transportation (denoted by v) and the time cost of travelling (denoted by t). The variable cost varies with how far the households stays away from the CBD and similarly the time cost accounts for income the households is giving up because of staying away from the CBD and travelling to work. We make some more assumptions on the travelling costs because each transportation mode has its own benefits and costs.

Assumptions on the transportation costs:

- $f^a > f^b$
- $v^a < v^b$
- $t^a < t^b$

These assumptions basically mean that travelling by bus has a higher fixed cost in comparison to bus, but time and variable cost is less if one chooses car as opposed to bus. Comparing the costs involved in the two transportation modes, a household choose with mode of transportation to use. A household staying at a distance x from the CBD chooses bus over car if and only if the following is true

$$f^b + v^b x + w_i t^b x < f^a + v^a x + w_i t^a x$$

$$\rightarrow x < \frac{f^a - f^b}{(v^b - v^a) + (t^b - t^a) w_i} \text{ for } i = H, L .$$

Therefore, under the assumption that travelling by bus involves more time and variable cost and the fixed cost of travelling by car is higher than the bus, we will find $\hat{x}_i$ such that households staying at a distance $\hat{x}_i$ from the CBD are indifferent between the two modes of transportation. This implies that if at a certain distance x from the CBD, the rich prefer bus, then the poor will also prefer bus at that x. This happens because the income of the rich (w_H) is greater than the income of the poor (w_L). Since we have assumed that the rich consume high quality of housing (q_H) and the poor consumes low quality of housing (q_L), the household will try to maximize its consumption of non-housing good (c). As has been assumed in Brock and Wrede (2008), we assume that housing consumption does not enter the utility function. So, the utility of a household living at a distance x from the CBD has utility

$$c_i^j(x) = w_i - f_j - v^j x - w_i t^j x - r_i^j(x) q_i \qquad (1),$$

for $i = H, L$ and $j = a, b$. $r_i^j(x)$ denote the rent at a distance x from the CBD. Since we have an open city model, we can assume that each type of household will attain the same utility level, so we can equate equation (1) to the exogenously given indirect utilities to get the bid-rent functions. The bid-rent functions would be as follows

$$r_i^j(x) = (w_i - f^j - v^j x - w_i t^j x - V_i)/\, q_i. \qquad (2)$$

In this section, we will modify the utility of the rich households to incorporate for the negative externality they get because of the existence of the low quality housing. Let Q be the total low quality housing in the city. This to also the total number of poor households in the city since we have assumed that only the poor households consume low quality housing (q_L). For simplicity, we assume that the disutility due to the low quality housing in the city enters linearly in the utility function of the rich households, which implies the utility of the rich household living at a distance x from the CBD is

$$c_H^j(x) - \theta Q = w_H - f^j - v^j x - w_H t^j x - r_H^j(x) q_H - \theta Q \text{ for } j = a, b.$$

$\theta > 0$ captures the intensity of the disutility of having low quality housing in the city. Hence the bid-rent functions of the rich households are

$$r_H^j(x) = (w_H - f_j - v^j x - w_H t^j x - \theta Q \; - V_H)/\, q_H.$$

Now, we will take a step back, and work with the bid-rent functions when there is no negative externality and look at the different situations that can occur. Note that from our discussion on the choice of transportation mode, we can get the bid-rent function for each type of household. We have shown that we can find $\hat{x}_i$ for $i = H, L$, such that households staying at a distance $\hat{x}_i$ from the CBD are indifferent between the two modes of transportation. The following diagram shows how for each type of household we get the bid-rent function across different distances (x) from the CBD.

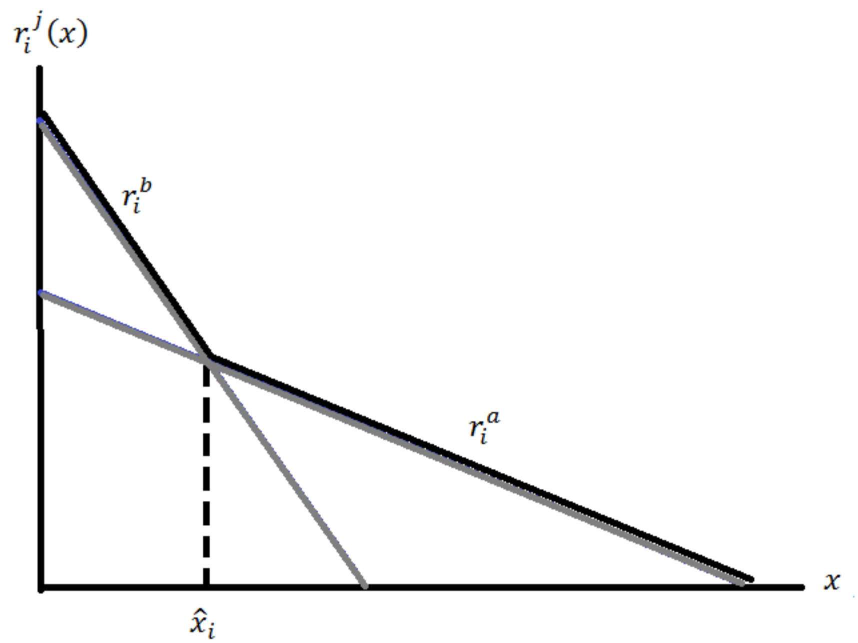


Figure 3.1

We see that after $\hat{x}_i$ the slope of the bid-rent function for either type of household changes because of the change in choice of transportation modes. Also, as mentioned before, since $\hat{x}_i = \frac{f^a - f^b}{(v^b - v^a) + (t^b - t^a)w_i}$ for $i = H, L$, $\hat{x}_H < \hat{x}_L$, because the rich households have higher income. When we bring the bid-rent functions of the two types of households together, we can get a complete picture of the different situations that can arise. Just for the sake of discussion, let us temporarily assume that land is only used for building houses and therefore housing developers are willing to produce housing if the bid-rent is positive. We will relax this assumption later, when we solve for the equilibrium. Also, note that unlike Brock and Wrede (2008), the bid-rent functions from the demand side would not tell us about the location pattern in the equilibrium because in our model we have two types of housing quality and we will model the housing supply side differently. Our discussion on the supply of housing will make it clear how our assumptions on the housing developers will affect location patterns in equilibrium.

However, it is important to discuss what would happen if follow Brock and Wrede (2008) and assume that whoever has the higher bid gets to stay at that place and housing of that quality is produced at that x. More formally, final rent at a distance x from the CBD is $r(x) = \max\{0, r_H(x), r_L(x)\}$.

Case 1: Poor live near the CBD, poor use bus and rich use car.
The following diagram depicts the situation when poor households live closer to the CBD, the poor households choose bus for transportation and the rich households use car for transportation.

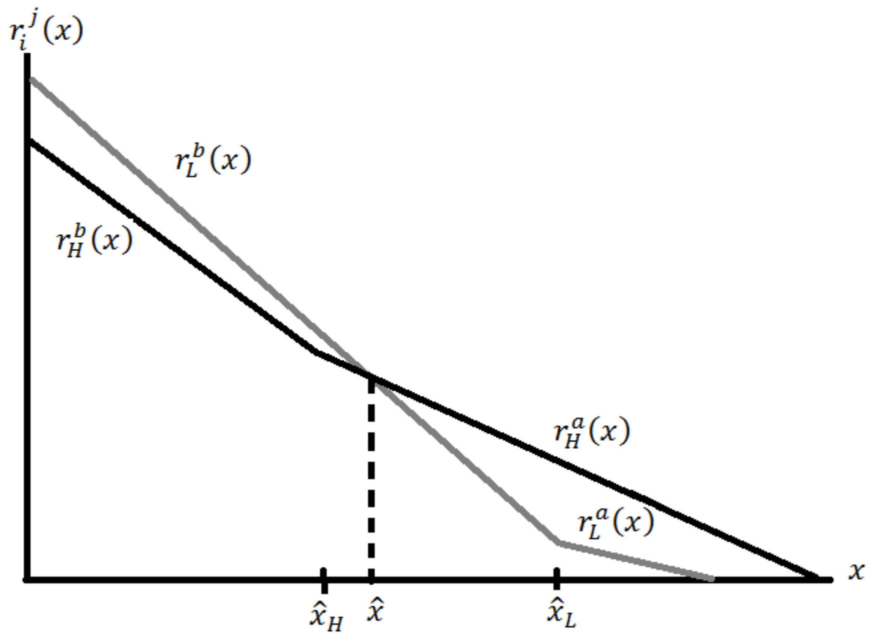


Figure 3.2

$\hat{x}$ in the diagram above is the point beyond which the bid-rent of the rich dominate that of the poor. The condition which characterizes $\hat{x}$ is as follows

$r_L^b(\hat{x}) = r_H^a(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^b + w_L t^b}{q_L} > \frac{v^a + w_H t^a}{q_H}$ which means that at the border between the poor and the rich, the poor household's bid-rent when they take the bus should be steeper than the rich household's bid-rent function when they take the car.
- $r_L^b(x = 0) > r_H^b(x = 0)$ which means that at distance zero from the CBD, the bid-rent of the poor household dominate that of the rich.

Case 2: Poor live near the CBD, poor use bus, rich use both bus and car

The following diagram depicts the situation where poor live closer to the CBD and use the bus for transportation. Beyond $\hat{x}$, upto $\hat{x}_H$, the rich households use the bus as a mode of transport and rich households living at a distance $x > \hat{x}_H$, use car for transportation.

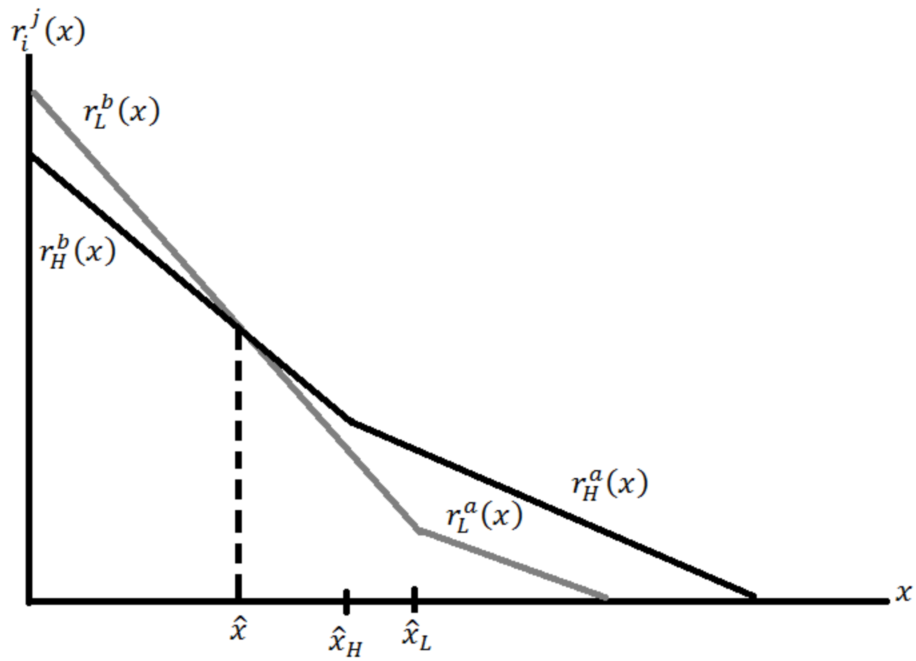


Figure 3.3

Similarly, as before, $\hat{x}$ in the diagram above is the point beyond which the bid-rent of the rich dominate that of the poor. The condition which characterizes $\hat{x}$ is as follows
$r_L^b(\hat{x}) = r_H^b(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^b + w_L t^b}{q_L} > \frac{v^b + w_H t^b}{q_H}$ which means that at the border between the poor and the rich, the poor household's bid-rent when they take the bus should be steeper than the rich household's bid-rent function when they take the bus.
- $r_L^b(x = 0) > r_H^b(x = 0)$ which means that at distance zero from the CBD, the bid-rent of the poor household dominate that of the rich.

Case 3: Poor live near the CBD, poor use bus and car, rich use car.

The following diagram depicts the situation where the poor live closer to the CBD and poor households living at a distance $x < \hat{x}_L$ use bus to commute to work. Poor households living at a distance x, where $\hat{x}_L < x < \hat{x}$, use car to commute to the CBD for work. Rich households living at a distance x, such that $x > \hat{x}$, use car to commute to the CBD for work.

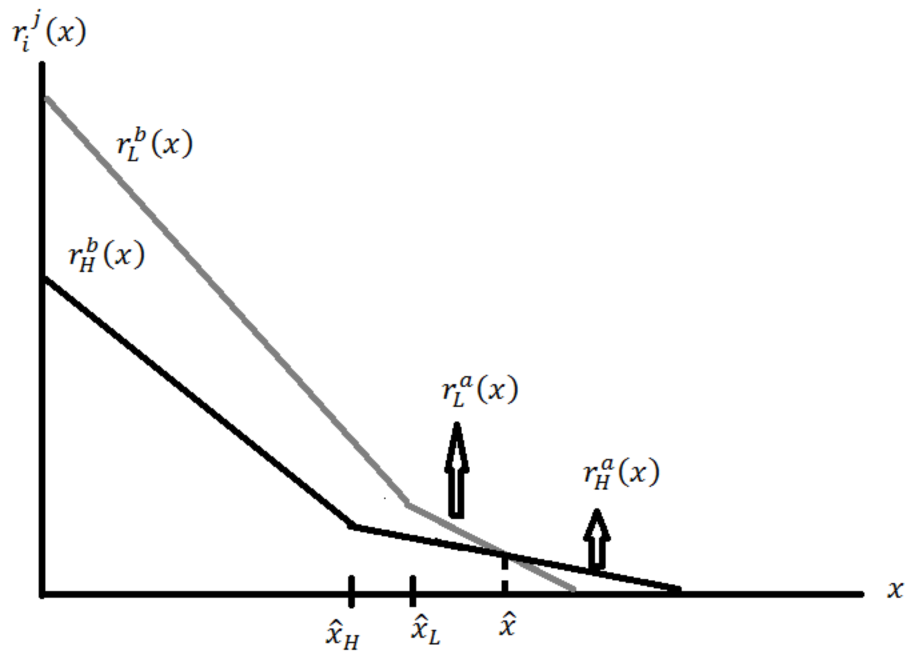


Figure 3.4

$\hat{x}$ in this diagram is the point beyond which the bid-rent of the rich dominate that of the poor. The condition which characterizes $\hat{x}$ is as follows

$r_L^a(\hat{x}) = r_H^a(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^a + w_L t^a}{q_L} > \frac{v^a + w_H t^a}{q_H}$ which means that at the border between the poor and the rich, the poor household's bid-rent when they take the car should be steeper than the rich household's bid-rent function when they take the car.
- $r_L^b(x = 0) > r_H^b(x = 0)$ which means that at distance zero from the CBD, the bid-rent of the poor household dominate that of the rich.

Case 4: Rich live near the CBD, rich use bus and poor use car.

The following diagram depicts the situation when rich households live closer to the CBD, the rich households choose bus for transportation and the poor households use car for transportation.

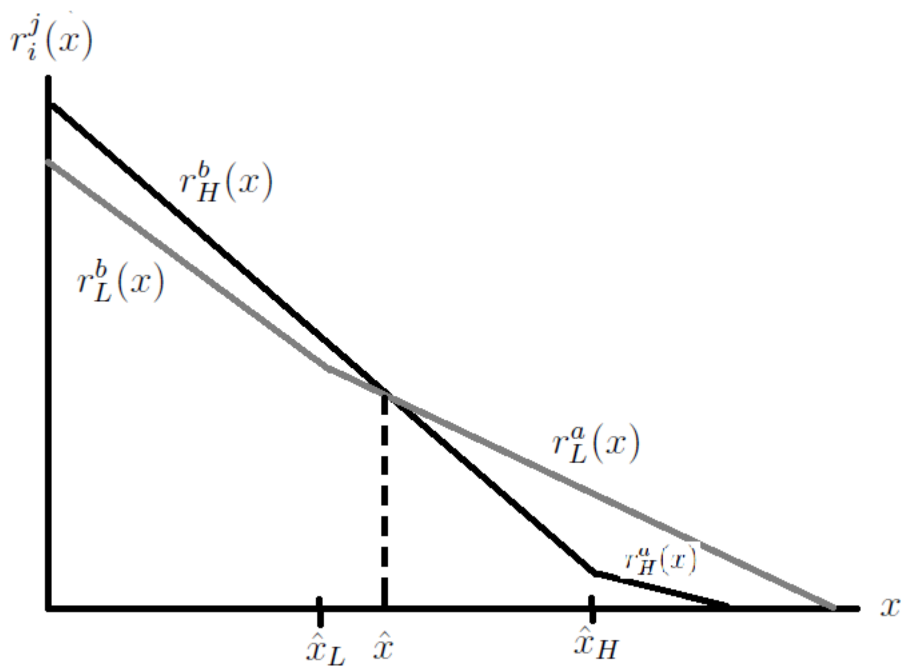


Figure 3.5

$\hat{x}$ in the diagram above is the point beyond which the bid-rent of the poor dominate that of the rich. The condition which characterizes $\hat{x}$ is as follows

$r_H^b(\hat{x}) = r_L^a(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^b+w_H t^b}{q_H} > \frac{v^a+w_L t^a}{q_L}$ which means that at the border between the poor and the rich, the rich household's bid-rent when they take the bus should be steeper than the poor household's bid-rent function when they take the car.

- $r_H^b(x=0) > r_L^b(x=0)$ which means that at distance zero from the CBD, the bid-rent of the rich household dominate that of the poor.

Case 5: Rich live near the CBD, rich use bus, poor use both bus and car

The following diagram depicts the situation where rich live closer to the CBD and use the bus for transportation. Beyond $\hat{x}$, upto $\hat{x}_L$, the poor households use the bus as a mode of transport and poor households living at a distance $x > \hat{x}_L$, use car for transportation.

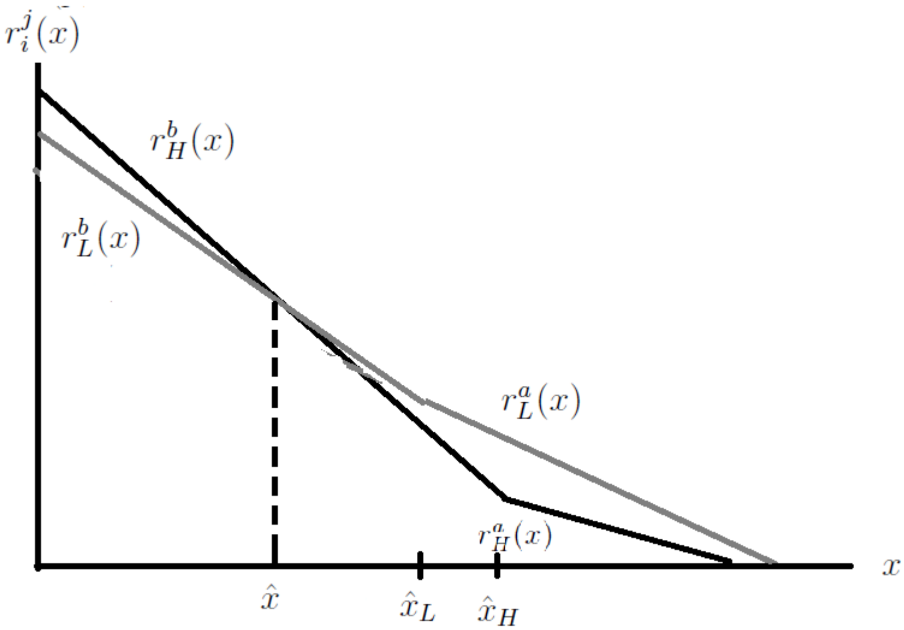


Figure 3.6

Similarly, as before, $\hat{x}$ in the diagram above is the point beyond which the bid-rent of the poor dominate that of the rich. The condition which characterizes $\hat{x}$ is as follows

$r_L^b(\hat{x}) = r_H^b(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^b + w_H t^b}{q_H} > \frac{v^b + w_L t^b}{q_L}$ which means that at the border between the poor and the rich, the rich household's bid-rent when they take the bus should be steeper than the poor household's bid-rent function when they take the bus.

- $r_H^b(x=0) > r_L^b(x=0)$ which means that at distance zero from the CBD, the bid-rent of the rich household dominate that of the poor.

Case 6: Rich live near the CBD, rich use bus and car, poor use car.

The following diagram depicts the situation where the rich live closer to the CBD and rich households living at a distance $x < \hat{x}_H$ use bus to commute to work. Rich households living at a distance x, where $\hat{x}_H < x < \hat{x}$, use car to commute to the CBD for work. Poor households living at a distance x, such that $x > \hat{x}$, use car to commute to the CBD for work.

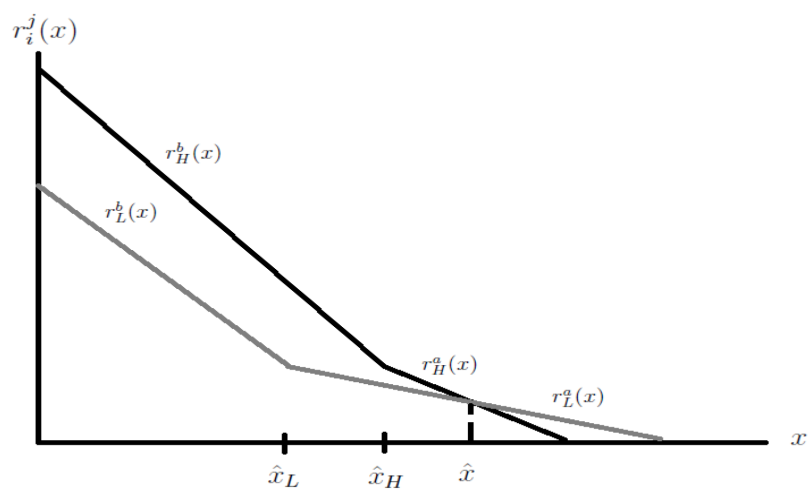


Figure 3.7

$\hat{x}$ in this diagram is the point beyond which the bid-rent of the poor dominate that of the rich. The condition which characterizes $\hat{x}$ is as follows

$r_H^a(\hat{x}) = r_L^a(\hat{x})$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^a+w_H t^a}{q_H} > \frac{v^a+w_L t^a}{q_L}$ which means that at the border between the poor and the rich, the rich household's bid-rent when they take the car should be steeper than the poor household's bid-rent function when they take the car.

- $r_H^b(x = 0) > r_L^b(x = 0)$ which means that at distance zero from the CBD, the bid-rent of the rich household dominate that of the poor.

Case 7: The following diagram shows a slightly more complicated scenario where the rich live closer to the CBD and take the bus. Then after a point ($\hat{x}_1$) poor households live and take the bus to travel but this changes at ($\hat{x}_2$), where the bid rent of the rich dominates that of the poor again and the rich use car for transportation.

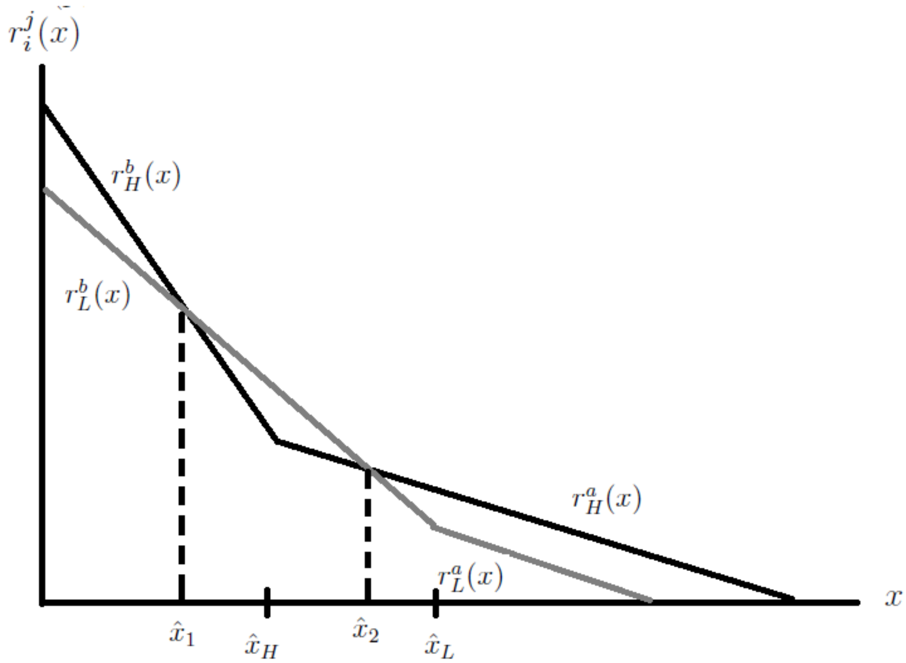


Figure 3.8

$\hat{x}_1$ and $\hat{x}_2$ are characterized by the following two equations:

- $r_H^b(\hat{x}_1) = r_L^b(\hat{x}_1)$,
- $r_H^a(\hat{x}_2) = r_L^b(\hat{x}_2)$

The conditions which needs to be satisfied such that we have the situation described in the diagrams are:

- $\frac{v^b+w_H t^b}{q_H} > \frac{v^b+w_L t^b}{q_L}$ which means that at the border (at $\hat{x}_1$) between the poor and the rich, the rich household's bid-rent when they take the bus should be steeper than the poor household's bid-rent function when they take the bus.
- $\frac{v^b+w_L t^b}{q_L} > \frac{v^a+w_H t^a}{q_H}$ which means that at the border (at $\hat{x}_2$) between the poor and the rich, the poor household's bid-rent when they take the bus should be steeper than the rich household's bid-rent function when they take the car.
- $r_H^b(x=0) > r_L^b(x=0)$ which means that at distance zero from the CBD, the bid-rent of the rich household dominate that of the poor.

3.3.2 Supply Side

On the supply side of the model, we have two types of housing suppliers, one specializes in the production of low quality housing and the other specializing in the production of high quality housing. Given, the assumptions of the demand side of the model, we know that the housing supplier which specializes in the production of high quality housing will only build houses for the rich and the housing supplier specializing in low quality housing will produce houses for the poor. Land is owned by absentee landlords and land is given to the housing supplier who gives higher returns to the landlord. The housing suppliers have a CRS production function and the profits of the housing suppliers at any distance x from the CBD are given by the following equation:

$$\pi_i(x) = r_i^j(x)h\big(s(x)\big) - c_i(x)s(x) - p_i^j(x),$$

where $i = \{H, L\}$ and $j = \{b, a\}$. Because of constant returns, housing output per unit of land can be expressed as $h(s(x))$, where h is the concave intensive form of the housing production per unit of land at distance x from the CBD, and $s(x)$ is the capital-to-land ratio at distance x from the CBD, referred to as structural density. $c_i(x)$ denotes the cost of building high or low quality housing at a distance x from the CBD. p_i denotes the return to the landlord. The housing supplier chooses $s(x)$ to maximize profits. The cost of production of the housing suppliers has two components: a fixed cost which does not vary with the distance and another component of the cost function varies with distance from the CBD. More specifically the cost function is:

$$c_i(x) = k_i + \alpha x$$

For $i = \{H, L\}$. We also assume that $k_H > k_L$, which implies that the fixed cost of producing high quality of housing is higher than the fixed cost of production of low quality housing. We also assume that $k_i > 0$. The housing suppliers choose $s(x)$ to maximize their profits. The optimality condition for the housing suppliers is as follows:

$$r_i^j(x)\frac{\partial h(s(x))}{\partial s(x)} = c_i(x) = k_i + \alpha x.$$

Since the housing suppliers earn zero profits, so the return to the landlord (p_i) is given by the following expression

$$p_i^j(x) = r_i^j(x)h\big(s(x)\big) - (k_i + \alpha x)s(x).$$

Note that, return to the landlord determines which type of housing supplier gets to build on a plot of land. From the previous analysis, it is not clear whether $p_H^j(x)$ dominates $p_L^j(x)$ or vice-versa.

To make the analysis more tractable, we will make some simplifying assumptions. In this model, we have assumed that the rich want good quality housing and the poor want low quality housing. Moreover, the quantity of housing is also fixed. Therefore, the housing supplier's problem boils down to simply choosing the number of households that can reside at a distance x from the CBD. Let us assume that, $h\big(s(x)\big) = N(x).\bar{q}_i$.

The housing suppliers choose $N(x)$ to maximize profits.

3.3.3 Equilibrium

Different Cost Function Specifications:

a) In this section, we will discuss the resulting equilibrium in the city when we incorporate the simplified version of $h\big(s(x)\big)$. Also, let the cost function be

$$c_i(x) = k_i + \alpha N(x).$$

The profit function of the supplier's is

$$\pi_i(x) = r_i^{\,j}(x) N_i(x)\bar{q}_i - \big(k_i + \alpha N_i(x)\big)\bar{q}_i \; - p_i^j(x).$$

One can compute the return to the landlord from the above expression and after simplifying we will get the following expression.

$$p_H^j(x) = [\big(r_H^j(x) - \alpha\big)N_H(x) - k_H]\bar{q}_H$$

and

$$p_L^j(x) = \big[\big(r_L^j(x) - \alpha\big)N_L(x) - k_L\big]\bar{q}_L.$$

Note that if $p_i^j(x) < 0$ for $i = \{H, L\}$ and $j = \{b, a\}$, then he would not be willing to rent to either of the housing suppliers (if we assume that rent from an alternative sector (say farming) yields zero returns). If at some x, for some value of $N_i(x)$, either $p_H^j(x)$ or $p_L^j(x)$ is positive, then land is given for building the type of housing that yields higher return. Since $p_i^j(x)$ is an increasing function of $N_i(x)$, if $p_i^j(x) > 0$ then the landlord would want to increase $N_i(x)$ as much as possible and this would lead to $N_i(x) \to \infty$. This cannot happen in equilibrium.

b) In this section, we will discuss the model under a different assumption on the cost of production. We would like to consider a situation where are cost of construction takes the following form:

$$c_i(x) = N_i(x)^2\bar{q}_i.$$

This implies that the profit function of housing suppliers would be given by,

$$\pi_i(x) = r_i^j(x)N_i(x)\bar{q}_i - N_i(x)^2\bar{q}_i - p_i^j(x)$$

The housing supplier chooses $N_i(x)$ to maximize profits. Since land is given to the supplier who gives the highest return to the landlord, so both the suppliers will be willing to give the maximum amount that they can.

Maximizing the above expression with respect to $N_i(x)$ gives us the optimal value of $N_i^*(x) = r_i^j(x)/2$. Therefore,

$$p_i^j(x) = \frac{1}{2}r_i^j(x)^2\bar{q}_i - \frac{1}{4}\bar{q}_i r_i^j(x)^2 = \bar{q}_i[\frac{1}{4}r_i^j(x)^2]$$

which can explicitly be written as

$$p_H^j(x) = \bar{q}_H\left[\frac{1}{4}r_H^j(x)^2\right],$$

and

$$p_L^j(x) = \bar{q}_L\left[\frac{1}{4}r_L^j(x)^2\right].$$

Note that, since $\bar{q}_H > \bar{q}_L$, if and if $r_H^j(x) > r_L^j(x)$, then we can say for sure that land would be given to the high type supplier. On the contrary, if $p_L^j(x) > p_H^j(x)$, then the low type housing supplier would get the land. The condition for this to happen is

$$\bar{q}_L\left[\frac{1}{4}r_L^j(x)^2\right] > \bar{q}_H\left[\frac{1}{4}r_H^j(x)^2\right]$$

$$\rightarrow \frac{r_L^j(x)^2}{r_H^j(x)^2} > \frac{\bar{q}_H}{\bar{q}_L}$$

Let us discuss two of the scenarios that we discussed before, specifically Cases 3 and 6. Note that from the previous analysis, we know that high quality housing supplier will get the land to build unless the squared ratio of the low and high quality housing is greater than the ratio of high to low quality housing.

Case 3: Recall the bid-rent function in case 3.

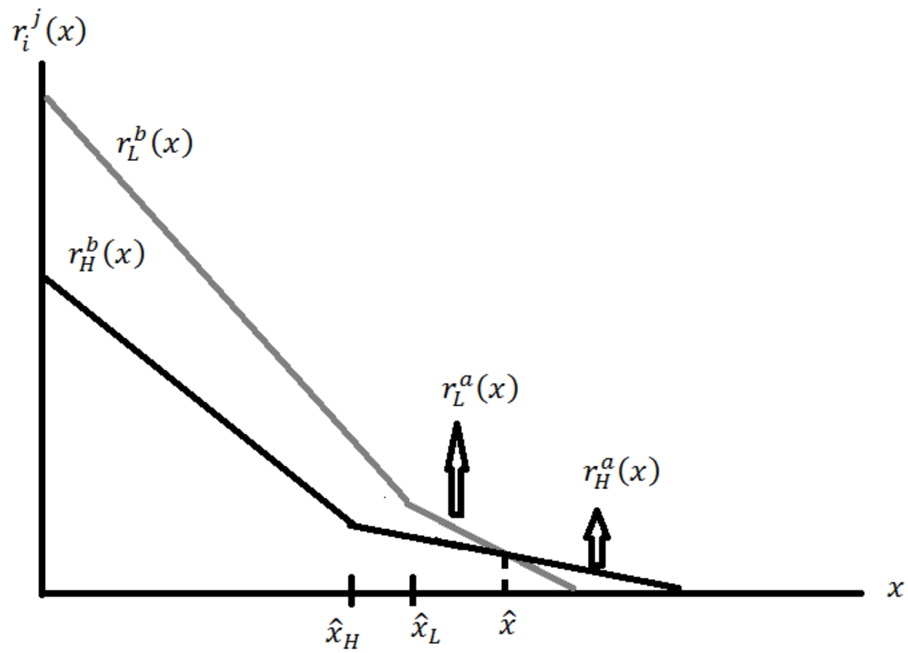


Figure 3.9

Note that the difference between the low and high quality bid-rents are highest at a distance of zero from the CBD and decreases thereafter. So, if

$$\frac{r_L^b(x=0)^2}{r_H^b(x=0)^2} > \frac{\bar{q}_H}{\bar{q}_L}$$

then, we will see low quality housing being built in the city. Otherwise, there would not be any low quality housing in the city. If the condition above holds, then there will exist some $\tilde{x},$ such that,

$$\frac{r_L^b(\tilde{x})^2}{r_H^b(\tilde{x})^2} = \frac{\bar{q}_H}{\bar{q}_L}.$$

So, in equilibrium, there will be low quality housing upto to distance $\tilde{x}$ from the CBD and beyond that there will be low quality housing. Note that, one can easily prove that $\tilde{x} \leq \hat{x}$.

In equilibrium, the number of poor and rich households the city can accommodate is as follows:

$$\int_0^{\tilde{x}} N_L^*(x) = N_P$$

$$\int_{\tilde{x}}^{\tilde{\tilde{x}}} N_H^*(x) = N_R,$$

where $\tilde{\tilde{x}}$ represents the point where the bid-rent function reaches zero (in this case, the bid-rent of the rich taking the car to work).

Proposition 3.1 *Under the situation as in Case 3, the city will be able to accommodate poor people only if* $\frac{r_L^b(x=0)^2}{r_H^b(x=0)^2} > \frac{\bar{q}_H}{\bar{q}_L}$, *otherwise there will be no poor people in the city.*

Case 6: Recall the bid-rent in case 6.

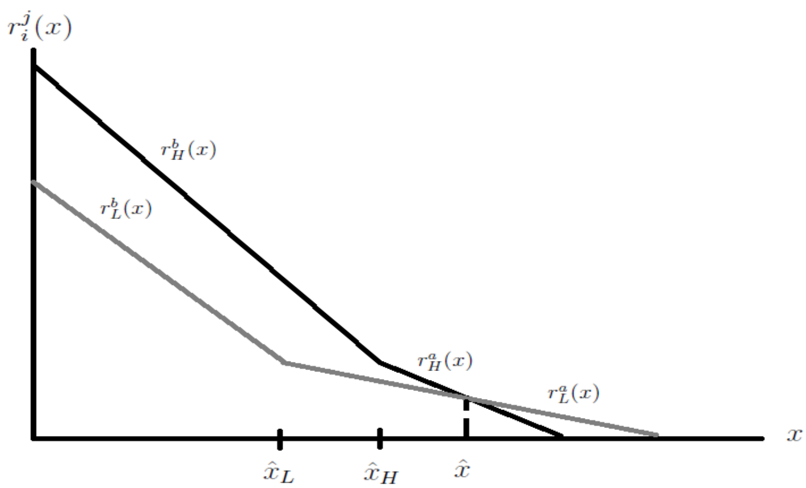


Figure 3.10

In this case, the high quality bid-rent dominates the low quality from the beginning, so upto $\hat{x}$, there will be high quality housing. Beyond $\hat{x}$, there is scope for low quality housing and that will be built only if the condition for low-quality building is met.
In this situation, we will have rich people living upto a distance of $\hat{x}$ for sure, but beyond $\hat{x}$, where the bid-rent of the poor dominate the rich, it might be the case that

high quality housing is build and the rich people stay. Once again let $\tilde{x}$ be such that $\frac{r_L^b(\tilde{x})^2}{r_H^b(\tilde{x})^2} = \frac{\bar{q}_H}{\bar{q}_L}$.

One can easily show that, $\tilde{x} > \hat{x}$. Also let us denote x_0, such that $r_H^a(x_0) = 0$. We can say for sure that beyond x_0, only low quality housing would be built and the poor will stay over there. Again, in equilibrium the number of rich and poor households the city can accommodate is as follows:

$$\int_0^{\tilde{x}} N_H^*(x) = N_R$$

$$\int_{\tilde{x}}^{\tilde{\tilde{x}}} N_H^*(x) = N_R,$$

Where $\tilde{\tilde{x}}$ represents the point where the bid-rent function reaches zero (in this case, the bid-rent of the rich taking the car to work).

Proposition 3.2 *Under the situation as in Case 3, the city will always have poor people in the city.*

c) In this case, we augment the cost of production further, and is

$$c_i(x) = N_i(x)^2 c_i \bar{q}_i$$

Therefore, the profit function of the housing suppliers would be

$$\pi_i(x) = r_i^j(x) N_i(x) \bar{q}_i - N_i(x)^2 c_i \bar{q}_i - p_i^j(x)$$

Using a similar procedure, to the case above we will have the returns to the landlord as follows:

$$p_H^j(x) = \frac{\bar{q}_H}{c_H}\left[\frac{1}{4} r_H^j(x)^2\right],$$

and

$$p_L^j(x) = \frac{\bar{q}_L}{c_L}\left[\frac{1}{4} r_L^j(x)^2\right]$$

Note that, with the new cost function specification in this section, low quality housing will be built only if the following condition is satisfied:

$$\frac{\bar{q}_L}{c_L}\left[\frac{1}{4}r_L^j(x)^2\right] \geq \frac{\bar{q}_H}{c_H}\left[\frac{1}{4}r_H^j(x)^2\right]$$

$$\rightarrow \frac{r_L^j(x)^2}{r_H^j(x)^2} \geq \frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}$$

Let us again recall cases 3 and 6 for further discussion. For the sake of simplicity in discussion, let us denote $\frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}$ as K. Note that $\frac{\bar{q}_H}{\bar{q}_L} > 1$ but $\frac{c_L}{c_H} < 1$, which implies that $K \geq 1$ or < 1.

i) $\boldsymbol{K > 1}$ Note that carefully looking at the bid-rent equation in Case 3, we can see that at $x = 0$, the poor outbids the rich but for the poor to stay there, the difference between the bid-rent functions should be such that

$$\frac{r_L^j(x)^2}{r_H^j(x)^2} \geq K.$$

Let us assume that $\frac{r_L^j(x=0)^2}{r_H^j(x=0)^2} \geq K$. Since the difference between the bid-rent functions are decreasing, there will be a point some $\tilde{x}$, such that

$$\frac{r_L^b(\tilde{x})^2}{r_H^b(\tilde{x})^2} = K.$$

In equilibrium, the number of poor and rich households the city can accommodate is as follows:

$$\int_0^{\tilde{x}} N_L^*(x) = N_P$$

$$\int_{\tilde{x}}^{\tilde{\tilde{x}}} N_H^*(x) = N_R,$$

Where $\tilde{\tilde{x}}$ represents the point where the bid-rent function reaches zero (in this case, the bid-rent of the rich taking the car to work).

ii) $\boldsymbol{K = 1}$ When $K = 1$, we see that the bid-rent functions dictate who lives where, that is the difference is quality of housing and the difference in construction cost of the two types of housing is such that they cancel each other's impact in the model and housing is provided to people who have the highest bid-rent at any x.

In equilibrium, the number of poor and rich households the city can accommodate is as follows:

$$\int_{0}^{\hat{x}} N_L^*(x) = N_P$$

$$\int_{\hat{x}}^{\tilde{\tilde{x}}} N_H^*(x) = N_R,$$

Where $\tilde{\tilde{x}}$ represents the point where the bid-rent function reaches zero (in this case, the bid-rent of the rich taking the car to work).

iii) $\boldsymbol{K < 1}$ In this situation, since $K < 1$, we can say for sure, that poor households will live in the city. Since the bid-rent of the poor is higher than that of the rich for all $x \leq \hat{x}$, we know that there will be a point some $\tilde{x}$, such that

$$\frac{r_L^a(\tilde{x})^2}{r_H^a(\tilde{x})^2} = K.$$

One can easily show that $\tilde{x} > \hat{x}$. Therefore in equilibrium, the number of poor and rich households the city can accommodate is as follows:

$$\int_{0}^{\tilde{x}} N_L^*(x) = N_P$$

$$\int_{\tilde{x}}^{\tilde{\tilde{x}}} N_H^*(x) = N_R,$$

where $\tilde{\tilde{x}}$ represents the point where the bid-rent function reaches zero (in this case, the bid-rent of the rich taking the car to work).

The observations in this section is summarized in the propositions below.

Proposition 3.3 *Augmenting the model, to include for construction cost heterogeneity, we find that results crucially depend on the ratio* $(\frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H})$. *Also, only when* $K > 1$, *we can have a situation when there are no poor households in the city. Similarly, only when* $K < 1$, *we can have a situation when there are no rich households in the city.*

Proposition 3.4 *K is crucial in determining the number of rich or poor households that will live in the city. Number of poor households in the city is directly proportional to K, and number of rich households in the city is inversely proportional to K.*

One can do a similar analysis for Case 6 as well, but it would be identical to the previous section, only the roles for the rich and the poor would be reversed.

3.4 Comparative Statics

In this section, we will extend the analysis by looking at how the outcome of the model changes if there are changes to the costs of construction. Let us continue the discussion using Case 3. Recall that low quality housing would be built if

$$\frac{r_L^j(x)^2}{r_H^j(x)^2} \geq \frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}.$$

We denoted $\frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}$ as K for simplicity in exposition. Let us consider the situation where c_L changes. As before, to start with, we can be in three different situations depending on the value of K.

a) This means that $\frac{\bar{q}_H}{\bar{q}_L}\frac{c_L}{c_H} > 1$, therefore an increase in c_L would imply that K would still be greater than 1. Let us also assume that $\frac{r_L^j(x=0)^2}{r_H^j(x=0)^2} \geq K$. Then, an increase in c_L would increase the value of K. Note that because of the increase in K, the distance at which $\frac{r_L^j(x)^2}{r_H^j(x)^2} < \frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}$ would move to the left to where it was before. As a result of this there would be less number of poor households in the city and more number of rich households in the city. On the other hand, if there is a decrease in c_L, then the resulting K after the change in the cost of construction can be greater, equal or less than 1. If K is still greater than 1, then the distance at which $\frac{r_L^j(x)^2}{r_H^j(x)^2} < \frac{\bar{q}_H}{\bar{q}_L} \cdot \frac{c_L}{c_H}$ would move to the right to where it was before. This would imply more poor households in the city and less rich households in the city. If, due to decrease in c_L, K becomes less than or equal to 1, then we move to the situations discussed in the previous section. It would still result in more poor households in the city and less rich households. The interesting point to note here is the fact that as long as $K > 1$, the point where there is a shift, that is the distance from the CBD where poor people stop residing and rich people start to live remains to the left of $\bar{x}$. If due to decrease in c_L, K moves from greater than one to less than one, then the shifting point moves from the left of $\bar{x}$ to the right of $\bar{x}$ or is equal to $\bar{x}$.

b) $K = 1$ When we start from a situation where $K = 1$, then a decrease in $c_L$would lead to a value of $K < 1$ and an increase in c_L would lead a value of $K > 1$. It is easy to see that in this situation an increase (decrease) in c_L would lead to less (more) number of poor households in the city.

c) $K < 1$ The analysis of this case would be very similar to the situation discussed in previously in this section, it would just be in the reverse direction.

Usually in the monocentric city models, the consumer with the highest bid-rent gets to stay over there. In our model, since the bid-rent alone does not decide who stays where, we have situations where the rich (poor) have a higher bid-rent but since high (low) quality housing is not provided, they are not able to stay there. As discussed before, the shifting point (let us denote this by $\tilde{\tilde{x}}$) can be to the left or right of $\hat{x}$ or $\tilde{\tilde{x}}$ can be equal to $\hat{x}$. As we have seen in the discussion in this section, changes in c_L

might lead to changes in position of $\tilde{\tilde{x}}$ with respect to $\hat{x}$. Let us call this phenomenon as regime shift.

Proposition 3.5 *Increase (decrease) in c_L would lead to a decrease (increase) in the number of poor households in the city and increase (decrease) in the number of rich households in the city. The results would go in opposite direction if there is an increase (decrease) in c_H.*

Proposition 3.6 *Increase (decrease) in c_L would lead to changes in the value of K which might further result in a regime shift.*

3.5 Conclusion

In this paper, we have taken used the monocentric city model and have enhanced the model to include two types of qualities of housing with different construction costs. Then we carefully depict all the different shapes the bid-rent functions can take depending on the parameter values of the model and show the city equilibrium patterns for all the different case in the absence of two types of housing suppliers. After that we introduce two types of housing suppliers and discuss a variety of construction cost specifications and discuss the equilibrium of the city. We move on to further enrich the model by introducing construction costs differences in building different types of housing and discuss two the many cases to show how the equilibrium changes when we include these additional features in the model. We show that in equilibrium, we might have situations where there are no poor households in the city or no rich households in the city. We also show that having the highest bid-rent does not ensure that that household will reside in that plot, it crucially depends on the supply side and it be the case that the preferred quality of housing is not supplied at that plot. We also show that results crucially depend on the ratio of high and low quality housing and the ratio of their respective construction costs. We also perform a comparative static analysis where we consider changes in the cost of construction of different types of housing. We find that changes in construction costs changes the number of rich and poor households the city can accommodate and moreover it might also cause changes in regimes. Overall, this paper contributes to the literature by incorporating new features in a monocentric city model and showing how these features contribute to the new equilibrium. One can also take this model as a base and further introduce construction cost differential across city space and further enrich the model.

3.6 References

1. Alonso, W. (1964) “Location and land use”, *Cambridge: Harvard University Press*.

2. Barnhardt, S., Field, E., & Pande, R. (2017) “Moving to Opportunity or Isolation? Network Effects of a Slum Relocation Program in India”, *American Economic Journal: Applied Economics*, Volume 9, pp. 1-32.

3. Bento, A. M., Takeuchi, A., & Cropper, M. (2008), “The Welfare Effects of Urban Land Use Policies on Slum Dwellers: The case of Mumbai, India”, *Journal of Urban Economics*, Volume 64, pp. 65-84.

4. Bertaud, A., & Brueckner, J. (2004), “Analyzing Building-Height Restrictions: Predicted Impacts, Welfare Costs, and a Case Study of Bangalore, India”, *Policy Research Working Paper 3290, World Bank.*

5. Brueckner, J. (2013), “Urban Squatting with Rent-Seeking Organizers”, *Regional Science and Urban Economics*, Volume 43, pp. 561-569.

6. Brueckner, J. A., & Selod, H. (2009), “A Theory of Urban Squatting and Land-Tenure Formalization in Developing Countries”, *American Economic Journal: Economic Policy*, Volume 1, pp. 28-51.

7. Hoff, K., & Sen, A. (2005), “Homeownership, Community Interactions, and Segregation”, *American Economic Review*, Volume 95, pp. 1167-1189.

8. Hoy, M. & Jimenez, E. (1991), “Squatters’ Rights and Urban Development: An Economic Perspective” *Economica*, Volume 58, pp. 79-92.

9. LeRoy, S., & Sonstelie, J. (1983), “Paradise Lost and Regained: Transportation Innovation, Income and Residential Location”, *Journal of Urban Economics*, Volume 13, pp. 67-89.

10. Mills, E.S. (1967), “An Aggregative Model of Resource Allocation in a Metropolitan Area”, *American Economic Review*, Volume 57, pp. 197-210.

11. Muth, R.F. (1969), “Cities and housing”, *Chicago: University of Chicago Press*.

12. Rosen, S. (1974), “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”, *Journal of Political Economy*, Volume 82, pp. 34-55.

13. Shah, N. (2014), “Squatting on Government Land”, *Journal of Regional Science*, Volume 54, pp. 114-136.

14. Turnbull, G. K., (2004) “Squatting, Eviction and Development”, *Regional Science and Urban Economics*, Volume 38, pp. 1-15.